

PROGRESS OF THEORETICAL PHYSICS

703

Founded by H. Yukawa in 1946

Board of Editors

| | |
|--------------------------|---------------------------|
| K. HUSIMI, <i>Osaka</i> | M. KOBAYASI, <i>Kyoto</i> |
| M. KOTANI, <i>Tokyo</i> | T. MUTO, <i>Tokyo</i> |
| S. SAKATA, <i>Nagoya</i> | S. TOMONAGA, <i>Tokyo</i> |
| H. YUKAWA, <i>Kyoto</i> | |

Volume 16

JULY—DECEMBER

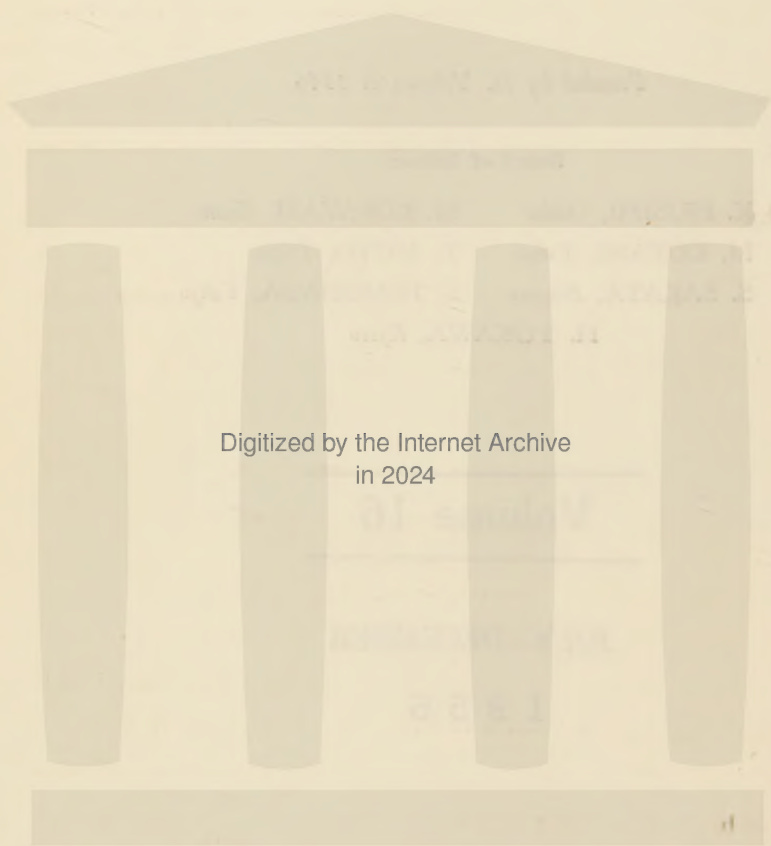
1 9 5 6

Published for

Research Institute for Fundamental Physics

The Physical Society of Japan

PROGRESS OF THEORETICAL PHYSICS



Digitized by the Internet Archive
in 2024

CONTENTS

No. 1. JULY

| | |
|---|------------------|
| A Classical Theory of the Collective Description | Yōiti Watanabe 1 |
| Nuclear Magnetic Relaxation in Antiferromagnetics | Tōru Moriya 23 |
| A Theory of Metallic Ferro- and Antiferromagnetism | |
| on Zener's Model | Tadao Kasuya 45 |
| Electrical Resistance of Ferromagnetic Metals | Tadao Kasuya 58 |

Letters to the Editor:

| | |
|---|--------------|
| The Decay Interaction of K_{u3} and K_{e3} and Fermi Interaction | 64 |
| S. Furuichi, T. Kodama, Y. Sugahara, M. Yonezawa and A. Wakasa | 64 |
| Interaction of K^+ -Meson and Nucleon | S. Minami 66 |
| Nature of Nuclear Forces indicated by the Photodisintegration of the Deuteron, II | 68 |
| S. H. Hsieh | 68 |

No. 2. AUGUST

| | |
|--|---------------------|
| On the Quantum Mechanics-like Description of the Theories of the Brownian Motion and Quantum Statistical Mechanics | 71 |
| Nobuhiko Saito and Mikio Namiki | 71 |
| Nucleon Magnetic Moments in Ps-Ps Intermediate Coupling Meson Theory | 95 |
| Akira Kanazawa and Masao Sugawara | 95 |
| On Quantum-Mechanical Nuclear Dipole Vibrations | Jun-ichi Fujita 112 |
| Energy Loss of a Charged Particle Traversing Superconductors | 131 |
| Satio Hayakawa and Kazuo Kitao | 131 |
| Energy Loss of a Charged Particle Traversing Ionized Gas and Injection Energies of Cosmic Rays.. Satio Hayakawa and Kazuo Kitao | 139 |

Letters to the Editor:

| | |
|--|--|
| An Attempt of Generalizing the Hypothesis of Charge Independence.... | O. Hara and Y. Fujii 149 |
| A Conjecture on the Possible Existence of New Hyperons..... | O. Hara and Y. Fujii 150 |
| Inelastic Scattering of Nucleons by Nuclei..... | R. Kajikawa, T. Sasakawa and W. Watari 152 |
| On the Theory of Thermal Conductivity of Pure Metals at Low Temperatures.. | P. G. Klemens 154 |

No. 3. SEPTEMBER

| | | |
|---|------------------------------------|-----|
| On the Quantum Field Theory of the Interaction between the Graviton and the Matter Field | Toshiei Kimura | 157 |
| Note on the Decay Interactions of Hyperons and Heavy Mesons | Chikashi Iso and Masaaki Kawaguchi | 177 |
| On the Determination of the Spins of Λ^0 and θ^0 Particles | Shoichi Hori | 189 |
| Partial Description of Quantum-Mechanical System | Gentaro Araki | 197 |
| Free-Electron Model of Benzene | Gentaro Araki | 212 |
| The Energy Band Structure of the Metallic Copper— <i>The Orthogonalized Plane Wave Method</i> — | Mitsuru Fukuchi | 222 |
| Extensions of Variational Methods, III— <i>Determination of Potential from Phase Shift Function</i> — | Takashi Ohmura | 231 |

Letters to the Editor :

| | | |
|--|---|-----|
| State Vectors in the Collective Description | T. Nishiyama | 244 |
| Some Remarks on the Nuclear Level Spacing | Y. Fujimoto, F. Fukuzawa and S. Okai | 246 |
| Note on the Decay of Ξ Particles | S. Hatano | 248 |
| On the Energy Dependence of the Cross Section for the Production of the Penetrating Shower Underground | S. Higashi, M. Oda, S. Oshio, H. Shibata, K. Watanabe and Y. Watase | 250 |
| On the Nuclear Interaction of μ -Meson below Ground | T. Kitamura and M. Oda | 252 |
| Thermodynamic Properties of He^3 | V. S. Nanda and S. K. Trikha | 254 |
| On the Energy Matrix of the Tensor Forces | T. Sasakawa | 256 |

No. 4. OCTOBER

| | | |
|---|---|-----|
| The Range of the Interaction between Different Fermions .. | Michiji Konuma | 259 |
| Free-Electron Ring Model of Polyacenes | Gentaro Araki | 265 |
| On the Grand Canonical Distribution Method of Statistical Mechanics | Tsunenobu Yamamoto and Hirotugu Matsuda | 269 |
| A Relativistic Wave Equation for a Particle with Two Mass States of Spin 1 and 0 | Haruo Shimazu | 287 |
| General Theory of Deuteron Induced Reaction, I — (d, p) and (d, n) Reactions — | Haruo Ui | 299 |

| | |
|--|--------------------|
| Nuclear Deformability and Shell Structure | |
| Toshio Marumori, Shôta Suekane and Atsuko Yamamoto | 320 |
| Nuclear Reactions in the Later Stage of the Stellar Evolution | |
| Kimiko Nakagawa, Takashi Ohmura, Hisao Takebe and Shinya Obi | 341 |
| On the Theory of Condensation | Kazuyosi Ikeda 361 |

Letters to the Editor :

| | | |
|--|-----------------------------|-----|
| A Lattice Model of Liquid Helium | T. Matsubara and H. Matsuda | 416 |
| On the Effects of Heavy Unstable Particles to the Anomalous Magnetic Moments of Nucleons | C. Iso and S. Iwao | 417 |

No. 5. NOVEMBER

| | | |
|---|---|-----|
| On a Solution of Field Equations in Einstein's Unified Field Theory, I | N. N. Ghosh | 421 |
| Deuteron Stripping and the Compound Nucleus | Masayuki Nagasaki | 429 |
| Theory of Bose-Einstein Condensation of an Imperfect Bose-Einstein Gas | Takeo Matsubara, Akira Morita and Namio Honda | 447 |
| Two-Nucleon Problem with Pion Theoretical Potential, I | | |
| — <i>Determination of Coupling Constant and Deuteron Problem</i> — | | |
| Junji Iwadare, Shoichiro Otsuki, Ryozyo Tamagaki and Wataro Watari | | 455 |
| Two-Nucleon Problem with Pion Theoretical Potential, II | | |
| — <i>Singlet Even State</i> — | | |
| Junji Iwadare, Shoichiro Otsuki, Ryozyo Tamagaki and Wataro Watari | | 472 |
| The Creation of an Electron Pair by a Fast Charged Particle | | |
| Toshiyuki Murota, Akira Ueda and Hajime Tanaka | | 482 |
| On the Foundation and the Applicability of Williams-Weizsäcker Method | Toshiyuki Murota and Akira Ueda | 497 |
| Helium Capturing Reactions in Stars | | |
| Satio Hayakawa, Chushiro Hayashi, Mitsuo Imoto and Ken Kikuchi | | 507 |

Letters to the Editor :

| | | |
|--|-----------------------|-----|
| Charge Independence Hypothesis and Leptons | T. Ouchi and K. Senba | 528 |
| Integral over Path Formulation of Statistical Theory of Irreversible Processes | D. Falkoff | 530 |

| | | |
|--|---------------------------|-----|
| On the Bethe-Salpeter Equation for the Scalar-Spinor Particle System | R. Sugano and Y. Munakata | 532 |
| Constraint in a Quantum System | Y. Watanabe | 534 |

No. 6. DECEMBER

| | | |
|---|--|------|
| On an Extended Frame work for the Description of Elementary Particles | Mineo Ikeda and Yoshihiko Miyachi | 537 |
| Meson-Nucleon Scattering in Low Energy Region | Daisuke Ito and Shigeo Minami | 548 |
| Quantum Field Theory of the Interaction between the Graviton and the Matter Field, II — <i>Formulation without the Coordinate Condition</i> — | Toshiei Kimura | 555 |
| A Lattice Model of Liquid Helium, I | Takeo Matsubara and Hiotsugu Matsuda | 569 |
| Some Consequences of the Compound Hypothesis for Elementary Particles | Ken-iti Matumoto | 583 |
| On the Bose-Einstein Condensation | Shigetoshi Katsura | 589 |
| Two-Nucleon Problem with Pion Theoretical Potential, III — <i>p-p Scattering at 18.2 Mev</i> — | Junji Iwadare, Shoichiro Otsuki, Ryoza Tamagaki and Wataro Watari | 604 |
| Formation of Light Nuclei in the Expanding Universe | Chushiro Hayashi and Minoru Nishida | 613 |
| The Composite Model for New Unstable Particles, I | Sho Tanaka | 625 |
| The Composite Model for New Unstable Particles, II | Sho Tanaka | 631 |
| Nuclear Magnetic Relaxation in Antiferromagnetics, II | Toru Moriya | 641 |
| Two-Nucleon Problem with Pion Theoretical Potential, IV — <i>Photodisintegration of Deuteron at 20 Mev</i> — | Junji Iwadare, Shoichiro Otsuki, Mitsuo Sano, Shuji Takagi and Wataro Watari | 658 |
| On a Theory of the Composite Model of Elementary Particles | Ziro Maki | 667 |
| Letters to the Editor : | | |
| A New Formulation on the Electromagnetic Field | T. Ohmura | 684 |
| Stability of the Electron | T. Ohmura | 685 |
| On a Composite Model for the New Particles | S. Sakata | 686 |
| Special Theory of Relativity and the Structure of Elementary Particles | H. Yukawa | 688 |
| CONTENTS AND AUTHORS INDEX TO VOLUME 16 | | i-vi |

A Classical Theory of the Collective Description

Yôiti WATANABE

Physics Department, Osaka University, Osaka

(Received March 12, 1956)

An elementary theory is developed of a classical treatment with the 'Collective Motion' of many-particle systems. The fundamental concept of collective motion can be reduced into a certain transformation property of Hamiltonian. Regarded as the dynamical variables, the continuous parameters of such a transformation describe the collective motion of the system, while the individual motions are partly suppressed by the same numbers of subsidiary conditions, or constraints, logically derived in, as the numbers of transformation parameters. Then the general procedure of formulation is illustrated; as an example, the incompressible motion with three modes can be systematically formulated. As another example, the unified treatment is performed also with the vortex motion. The digression in the approximating procedure is also presumed.

§ 1. Introduction

Recently, in treating various problems on the system of many particles, the notion of collective motion has awakened interest among many physicists.¹⁾⁻¹⁰⁾ Less attention is paid there to the motion of individual particles, but more stress is laid on the organized one of the whole system. D. Bohm and D. Pines have developed such a treatment for the problem of so-called plasma oscillation in a dense gas of electrons. They indicated that their organized behaviours with lower frequency could be transferred into the behaviours of electromagnetic field by means of a suitable canonical transformation, and that no attention would be consequently needed to the lower frequency part of the oscillations concerning the individual particles.¹²⁾ Their theory, however, apparently relying upon the long range nature of Coulomb forces, seems difficult to be applied to other important cases with the different nature of forces.

It is well-known that there is a sort of collective motion also in nuclear systems where the forces have no more long range nature. By the fundamental standpoint of Bohm's theory we can understand the possibility of transfers of part of the particle motion into the respondent motions of auxiliary quantities such as the electromagnetic field. This concept can be applied by an appropriate generalization in spite of the different nature of force range, e.g., to the problems of nuclear structure.

Tomonaga has investigated the similar problem independently from a quite different point of view¹³⁾. His method is a natural generalization of the separation of the translational motion from the relative motion with respect to the centre-of-mass. Corresponding to the c.m. co-ordinate, some co-ordinates have been introduced as derived from a sort of the velocity potential of irrotational fluid for a certain mode of collective motion. Thus

there are two stand-points: one, some auxiliary co-ordinates enter into the formalism as in Bohm's investigation, another, no such co-ordinates are used as in Tomonaga's work.

We want to follow the former stand-point by the general, classical way. The auxiliary co-ordinates describing the *classical* collective motions, (we do not here include such as spin wave), are introduced by means of continuous point transformations. For instance, the quantity to describe the c.m. motions is introduced as the parameters of space-transformation. This circumstance will be discussed in § 2. Similar procedure will appear in § 3 in the treatment of a simple incompressible motion in two dimensions. The results are generalised in § 4. In this treatment a number of new variables are introduced, and therefore the same number of subsidiary conditions have to be imposed among the variables to suppress the redundant degrees of freedom. They are automatically brought into the formalism by the generalized gauge condition of the modified Lagrange function. The relation to the hydrodynamics is investigated in § 5. We have the most manageable expressions in § 6 to take out the collective modes of the system. Then rather stringent constraints are put thereby instead of the prescribed subsidiary conditions by means of the *principle of stiffening*. Such variables formally introduced have, however, not always the physical significance. We have to take into account the statistical aspects, and postulate a requirement for the adiabatic approximation and so on; these are discussed in § 7. In the final results we have a number of Pfaffian constraints. Our theory may be related to Miyazima's theory of irrotational fluid model when these Pfaffians are integrable. Even though any of them are not integrable, a unified treatment is possible. In the hydrodynamical words, this corresponds to the vortex motion. These circumstances are illustrated by an example of rigid rotator in two dimensions; also a two-dimensional incompressible motion with three modes is analysed according to the general procedure in § 8.

We worked only in the classical domain because of lack of our knowledge for the quantization with respect to the quantities spanning a nonlinear topological space. In general, the collective variables introduced by such a way do not span the linear vector space,—even not the curvilinear one—, in which the usual quantization is available.

§ 2. Center of mass

The notion of centre-of-mass is one of the simplest and the most familiar ones of the collective motion in a many-particle system. Usually the c.m. motion is described by a co-ordinate canonically conjugate to the total linear momentum, but it may also be introduced by means of auxiliary variables. By way of illustration of our general procedure, we shall first explain how to introduce the notion of centre-of-mass for the case of N particles moving on a straight line in a seemingly roundabout manner which, however, is suitable for the generalization to more complicated modes.

(1) Let the Lagrangian of the particle system be

$$L_P \equiv (M/2) \sum \dot{\bar{x}}_i^2 - \bar{V}(\bar{x}_1 \cdots \bar{x}_N). \quad (2.1)$$

\bar{x}_i is the co-ordinate of the i -th particle, referred to a fixed origin, but we shall now refer

the system to a moving origin by

$$\bar{x}_i = x_i + \alpha. \quad (2.2)$$

In terms of new co-ordinates the Lagrangian becomes

$$L_I \equiv (M/2) \sum_i (\dot{x}_i + \dot{\alpha})^2 - \bar{V}(x_1 + \alpha, \dots, x_N + \alpha). \quad (2.3)$$

α may be used to represent a prescribed motion, but it may also be regarded as a redundant dynamical variable independent of the particle co-ordinates x_i . In the latter interpretation the canonical conjugates of x_i and α are interrelated by an identity

$$\sum_i \partial L_I / \partial \dot{x}_i - \partial L_I / \partial \dot{\alpha} = 0, \quad (2.4)$$

which prevents a straight-forward transformation to the canonical system. As has been shown by Rosenfeld in the similar case of electrodynamics, we may nevertheless proceed to set up a canonical system by introducing the canonical conjugates p_i of x_i and β of α , and by regarding the above-mentioned identity as a subsidiary condition for momenta:⁽¹⁴⁾⁻⁽¹⁶⁾

$$\phi \equiv \sum p_i - \beta = 0. \quad (2.5)$$

The Hamiltonian is determined up to a term $v\phi$ containing a completely arbitrary factor v :

$$H_I \equiv (1/2M) \sum p_i^2 + V(x_1 \dots x_N; \alpha) - v\phi. \quad (2.6)$$

We get from this Hamiltonian as equations of motion

$$dx_i/dt = p_i/M - v. \quad (2.7)$$

They contain an arbitrary velocity v . Summing them up, we get

$$\sum dx_i/dt = \sum p_i/M - Nv = \beta/M - Nv. \quad (2.8)$$

Now we want to specify our arbitrary v to be equal to β/NM , so that the sum of particle velocities vanishes,

$$\sum dx_i/dt = 0; \quad (\sum x_i = \text{const.}), \quad (2.9)$$

by which relation the variable α is identified with the c.m. co-ordinate. By this choice our Hamiltonian becomes

$$H_I \equiv (1/2M) \sum (p_i - \beta/N)^2 + \beta^2/2NM + \bar{V}(x + \alpha) \\ \text{with a subsidiary condition:} \quad \phi \equiv \sum p_i - \beta = 0. \quad (2.10)$$

Now let us set up the generating function of a second canonical transformation

$$W = \{ \sum x_i/N + \alpha \} B + \sum x_i p_i, \quad (2.11)$$

it transforms the Hamiltonian into

$$H_{II} \equiv (1/2M) \sum P_i^2 + B^2/2NM + \bar{V}(X_i + A - (1/N) \sum X_i) \\ \text{with a subsidiary condition:} \quad \phi \equiv \sum P_i = 0, \quad (2.12)$$

in terms of new variables (X, P, A, B) . Then a differential constraint

$$d\chi \equiv \sum dX_i = 0 \quad (2.13)$$

results from the equation of motion and the subsidiary condition. It is integrated to give the constant of motion

$$\sum X_i = NR_0 = \text{const.} \quad (2.14)$$

(2) We may stop here. The consideration of more complicated cases motivates, however, a further transformation. We have suppressed the excessive degree of freedom introduced as the collective variables α, β by means of *subsidiary* condition $\phi=0$ (2.13). By a subsidiary condition $\phi=0$ we understand a relation conserved in the temporal development of the system: that is, such a relation $\phi=0$ that $[H, \phi]=0$. Then it is sufficient to give the quantity the value ' $\phi=0$ ' at *one* epoch. At least in the classical mechanics (possibly also in the quantum mechanics) there is another way of suppressing excessive degrees of freedom. This is a kinematical condition, a *constraint* as opposed to the above dynamical condition. We can anticipate this kinematical condition in the form that the "centre-of-mass" co-ordinate $\sum X_i$ of the internal individual motion is at rest,—is fixed at the origin, say.

We have seen in the above that the kinematical $\sum \dot{X}_i = 0$ is a consequence of the dynamical condition $\sum P_i = 0$ if the Hamiltonian H_{II} is given by (2.12). Now we want to subject our system to the kinematical constraint $\sum dX_i = 0$. Generally speaking, this requires an additional, passive force of constraint to be included in the Hamiltonian $H \rightarrow H^* = H_{II} + u \sum X_i$, where u is a measure of that passive force. The resulting equations of motion, however,

$$\begin{aligned} \sum dX_i/dt &= [H^*, X_i] = \sum P_i/M, \\ \sum dP_i/dt &= [H^*, P_i] = -Nu \end{aligned} \quad (2.15)$$

make u vanish together with the imposed constraint $\sum X_i = NR_0$. It will be seen also that $\sum P_i = 0$ is a consequence of $\sum X_i = NR_0$ (const.). After we have confirmed that our mechanical system is equivalent to the system with the Hamiltonian (2.12) and is subject to the constraint (2.14)*, we try our final transformation in order to get a manageable form of the Hamiltonian. Such a transformation consists, so to speak, in getting back the original variable α, β instead of A, B .

(3) Let

$$W = (\alpha + R_0)B \quad (2.16)$$

be the generating function of the transformation $(A, B) \rightarrow (\alpha, \beta)$. Our Hamiltonian becomes

$$H_c \equiv (1/2M) \sum P_i^2 + \beta^2/2NM + \bar{V}(X_i + \alpha), \quad (2.17)$$

since we are always holding $\sum X_i/N$ fixed to be equal to R_0 .

* We know this verification as 'principle of stiffening' in the elementary mechanics

(4) This universal postulate of the constraint also applies to the definition of the Poisson brackets ; (P.b.), because the operation of differentiation which is to be replaced by

$$\partial/\partial X_i^* = \partial/\partial X_i - (1/N) \sum \partial/\partial X_k \quad (2.18)$$

is concerned with it. We shall, however, omit the detailed explanation with respect to this reduction of redundant degrees of freedom from the Hamiltonian system ; it will be done in a more systematic way in Appendix II.

(5) So far we did not say anything about the form of potential energy. The virtue of notion of centre-of-mass depends of course on the fundamental functional form of it, $V(X, \alpha)$. The notion becomes favourable, (i) if $V(x)$ is translationally invariant : i.e.,

$$V_I(x_i + \alpha) = V_I(x_i), \quad (2.19)$$

(ii) if, besides this invariant 'internal' potential, it contains symmetric linear terms (uniform field of gravitation) :

$$V_{II}(x_i + \alpha) = g \sum x_i + N\bar{g}\alpha, \quad (2.20)$$

and (iii) if, besides above-mentioned two types of potentials, there is another potential energy expressed as a symmetric sum of the square of the co-ordinates :

$$V_{III}(x_i + \alpha) = (\kappa/2) \sum (x_i + \alpha)^2. \quad (2.12)$$

(In the case where the masses of the particles are different from each other, the symmetric sum of (ii) & (iii) must be replaced by the mass-weighted sum.) If the particles are subject to external forces other than those of (ii) & (iii), the notion of center-of-mass becomes more or less approximate. Later we shall be concerned with the point a little. (See § 7).

In retrospect we see that the centre-of-mass has been introduced by subjecting each and every particle the same and arbitrary translation α , and by treating this α as a new degree of freedom of the system. This suggests that a large class of collective motion would be described by a continuous group of transformations on the variables of particles.

§ 3. Two dimensional incompressible and irrotational flow of particles

Let us investigate the problem of the so-called incompressible fluid model before proceeding to the general theory. For simplicity we shall operate in two dimensions though the three-dimensional case would be more important for nuclear physicists in regard to Bohr's ellipsoidal oscillations of nuclei. A two-dimensional incompressible motion will be considered by the presumed corresponding linear unimodular transformation :

$$\begin{aligned} \bar{x} &= (\alpha_0 + \alpha_3)x + (\alpha_1 - \alpha_2)y, & \bar{y} &= (\alpha_1 + \alpha_2)x + (\alpha_0 - \alpha_3)y, \\ \alpha_0^2 - \alpha_1^2 + \alpha_2^2 - \alpha_3^2 &= 1. \end{aligned} \quad (3.1)$$

The parameters α_1 and α_3 give rise to such deformations of $x-y$ plane that a circle with the center at the origin turns into two types of ellipses. The angle between their principal axes are equal to 45° . On the other hand α_2 brings about a plane rotation. We first

consider the simple case of shearing deformation $\alpha_1 = \alpha_2 = 0$, $\alpha_3 = \alpha \neq 0$:

$$\bar{x}_i = (\alpha_0 + \alpha)x_i, \quad \bar{y}_i = (\alpha_0 - \alpha)y_i, \quad (3.1)'$$

with $\alpha_0^2 - \alpha^2 = 1$.

The original Hamiltonian, expressed in terms of barred variables

$$H_p \equiv (1/2M) \sum (\bar{p}_x^2 + \bar{p}_y^2) + \bar{V}(\bar{x}_1, \dots, \bar{x}_N) \quad (3.2)$$

becomes

$$H_{I_0} \equiv (1/2M) \sum \{(\alpha_0 - \alpha)^2 p_x^2 + (\alpha_0 + \alpha)^2 p_y^2\} + V(x_1, \dots, x_N; \alpha),$$

with a subsidiary condition:

$$\phi \equiv \sum (x_i p_x^i - y_i p_y^i) - \alpha_0 \bar{J} = 0. \quad (3.2)'$$

Also this time the general form of the Hamiltonian

$$H_I \equiv H_{I_0} - v\phi \quad (3.3)$$

contains completely arbitrary factor v . By a second canonical transformation the variable β is banned out of ϕ . We put its generating function to be

$$W = A(x, y; \alpha) + B \sum (x_i p_x^i + y_i p_y^i), \quad (3.4)$$

where $A(x, \alpha)$ satisfies the differential equation

$$[\sum \{x_i \partial / \partial x_i - y_i \partial / \partial y_i\} - \alpha_0 \partial / \partial \alpha] A = 0. \quad (3.5)$$

It requires that A should never depend on x 's, y 's and α but only through \bar{x} 's and \bar{y} 's. The new variables (X, Y, P_x, P_y, A, B) can then be expressed in terms of the old ones ($x, y, p_x, p_y, \alpha, \beta$) through the expressions

$$\begin{aligned} X_i &= x_i, & Y_i &= y_i, & A &= A(x, y; \alpha), \\ p_x^i &= P_x^i + B \partial A / \partial x_i, & p_y^i &= P_y^i + B \partial A / \partial y_i, & \beta &= B \partial A / \partial \alpha, \end{aligned} \quad (3.6)$$

by which also the Hamiltonian is transformed into

$$\begin{aligned} H &\equiv (1/2M) \sum_i \{P_x^i (\alpha_0 - \alpha)^2 + P_y^i (\alpha_0 + \alpha)^2\} \\ &+ (1/2M) \sum_i \{(\alpha_0 - \alpha)^2 (\partial A / \partial x_i)^2 + (\alpha_0 + \alpha)^2 (\partial A / \partial y_i)^2\} \\ &+ \tilde{V}(X, Y; A) + (B/M) \sum_i \{(A \partial / \partial x_i) P_x^i (\alpha_0 - \alpha)^2 \\ &+ (\partial A / \partial y_i) P_y^i (\alpha_0 + \alpha)^2\} - v\tilde{\phi}, \end{aligned} \quad (3.7)$$

together with the subsidiary condition to be free from B by virtue of (3.5):

$$\tilde{\phi} \equiv \sum (X_i P_x^i - Y_i P_y^i) = 0.$$

In the above expression any of old variables x, y and α are interpreted to be expressed by new ones X, Y and A . We now choose the concrete form of $A(x, y, \alpha)$

$$A(x, y, \alpha) \equiv (1/2N) \sum \{(\alpha_0 + \alpha)^2 x_i^2 - (\alpha_0 - \alpha)^2 y_i^2\}, \quad (3.8)$$

so that the third term of (3.7) may be proportional to ϕ . The last two terms of (3.7) can amalgamate with each other, consequently. The second term becomes of the form $(1/2MN) \Gamma' B^2$ where $\Gamma' \equiv \sum \{(\alpha_0 + \alpha)^2 x_i^2 + (\alpha_0 - \alpha)^2 y_i^2\}$. Finally the form of the Hamiltonian is given by

$$H \equiv (1/2M) \sum \{P_x^2 (\alpha_0 - \alpha)^2 + P_y^2 (\alpha_0 + \alpha)^2\} + (1/2MN) \Gamma' B^2 + \tilde{V} - v\tilde{\phi}, \quad (3.7)'$$

where α 's are to be expressed in terms of X 's, Y 's and A .

The relation $[H_{II}, \tilde{\phi}] = 0$ can also be easily verified. In the same manner as before we can substitute such a subsidiary condition as suppressing the extra degrees of freedom alien to the internal motion by another kinematical condition

$$\alpha \tilde{\chi} \equiv \sum_i \{(\alpha_0 + \alpha)^2 X_i dX_i - (\alpha_0 - \alpha)^2 Y_i dY_i\} = 0, \quad (3.8)$$

or, expressed in terms of unbarred variables,

$$d\chi \equiv \sum_i \{(\alpha_0 + \alpha)^2 x_i dx_i - (\alpha_0 - \alpha)^2 y_i dy_i\} = 0, \quad (3.8)'$$

which alters nothing of the equations of motion. This time, however, we cannot find the integrated form of the constraint such as (2.14).^{*} Then $\phi = 0$ becomes a result of (3.8).

After such a stiffening, we try the final 'return transformation' as in the c.m. case so as to obtain a manageable form of Hamiltonian. Let

$$W = (1/2N) \sum \{(\alpha_0 + \alpha)^2 x_i^2 - (\alpha_0 - \alpha)^2 y_i^2\} B \quad (3.9)$$

be the generating function of the transformation. By virtue of the constraint the variables x_i , y_i must be considered as parameters in (3.9). Our Hamiltonian becomes then,

$$H_c \equiv (1/2M) \sum \{P_x^2 (\alpha_0 - \alpha)^2 + P_y^2 (\alpha_0 + \alpha)^2\} + (1/2M \Gamma') (\alpha_0 \beta)^2 + V(X, Y, \alpha). \quad (3.10)$$

Also in this case the constraint applies to the definition of differentiation operations. They are replaced as follows:

$$\begin{aligned} \partial/\partial X_i &\rightarrow \partial/\partial X_i^* = \partial/\partial X_i - (1/\Gamma') (\alpha_0 + \alpha)^2 X_i \sum (X_i \partial/\partial X_i - Y_i \partial/\partial Y_i), \\ \partial/\partial Y_i &\rightarrow \partial/\partial Y_i^* = \partial/\partial Y_i + (1/\Gamma') (\alpha_0 - \alpha)^2 Y_i \sum (X_i \partial/\partial X_i - Y_i \partial/\partial Y_i). \end{aligned} \quad (3.11)$$

In contrast to the c.m. case, the collective and the individual terms are not separated in the Hamiltonian (3.12). The first term, 'The particle kinetic energy term', contains the collective variables α , and the second, 'The collective kinetic energy term', contains the individual particle variables X 's and Y 's. More advanced considerations will be required for the practical separation of these terms. It will be discussed later in § 7.

§ 4. Lagrangian and Hamiltonian

Now we shall generalize the treatment in the preceding sections by supposing a

^{*} See also Appendix II.

continuous transformation with r parameters to correspond to r modes of a certain collective motion. In the beginning we have the Hamiltonian of a many-particle system in terms of barred variables \bar{x} and \bar{p} , called hereafter "The particle representation" or (P) in short; after the first transformation, we have the Hamiltonian in terms of unbarred variables x and p , called "The first representation" or (I).

The point transformation from (P) to (I) is given as

$$\bar{x}_i^\mu = S^\mu(x_i; \alpha^\sigma), \quad (4.1)$$

where \bar{x}_i^μ and x_i^μ are the co-ordinates of individual particles in (P) and (I) respectively; $\mu=1, 2, 3$ the indices for the components of vectors, $i=1 \cdots N$ the names of the constituent particles, and $\sigma=1, 2, \cdots r$ the indices of the transformation parameters. There is a set of differential equations

$$A_\sigma^\mu(x) \partial S^\nu / \partial x^\mu - G_\sigma^\nu(\alpha) \partial S^\nu / \partial \alpha^\sigma = 0 \quad (4.2)$$

for $S^\nu(x, \alpha)$, where A or G are determined by the group properties of the transformation.²⁰ Moreover the assumption

$$x^\mu = S^\mu(x, 0) \quad (4.3)$$

spoils nothing of the generality. (4.2) makes $\partial \bar{x}^\mu = 0$ when x 's and α 's are slightly displaced by

$$\delta x^\mu = A_\rho^\mu(x) \xi^\rho, \quad \delta \alpha^\sigma = G_\rho^\sigma(\alpha) \xi^\rho, \quad (4.4)$$

where ξ^ρ are arbitrary infinitesimal functions of time.

The Lagrangian L of the many-particle system is defined as the function of barred variables and their velocities; i.e., in terms of (P) variables,

$$L_p \equiv L_p(\bar{x}, \dot{\bar{x}}) \quad (4.5)$$

After the transformation (4.1) becomes

$$L_I \equiv L_I(x, \dot{x}; \alpha, \dot{\alpha}) \quad (4.6)$$

in terms of (I) variables. It is invariant under (4.4), quite analogously to the gauge invariance in the theory of electrodynamics. By virtue of such a 'gauge condition' we have r identities among the equations of motion

$$\sum_i A_\rho^\mu(x_i) \partial \bar{L}_I / \partial \dot{x}_i^\mu - G_\rho^\sigma(\alpha) \partial \bar{L}_I / \partial \dot{\alpha}^\sigma = 0, \quad (4.7)$$

where $\partial \bar{L} / \partial q = \partial \bar{L} / \partial q - d(\partial \bar{L} / \partial \dot{q}) / dt$ (Eulerian derivatives), as well as those among the canonical conjugates of x 's and α 's

$$\sum_i A_\rho^\mu(x_i) \partial L_I / \partial \dot{x}_i^\mu - G_\rho^\sigma(\alpha) \partial L_I / \partial \dot{\alpha}^\sigma = 0. \quad (4.7)'$$

They make impossible the 'unique' Legendre transformation to the canonical system. In order to set up a canonical system by introducing the canonical momenta p_μ' of x_i^μ and β_σ of α^σ the identities (4.7)' should be regarded as r linear subsidiary conditions

$$\sum_i A_p^\mu(\mathbf{x}_i) \dot{p}_\mu^i - G_p^\sigma(\alpha) \beta_\sigma = 0. \quad (4.8)$$

Also the Hamiltonian is obtained definitely up to the terms containing r indefinite multipliers v^ρ

$$H_I \equiv H_{I_0} - v^\rho \phi_\rho. \quad (4.9)$$

H_{I_0} is a special form of Hamiltonian derived from (4.6). The equations of motion

$$d\mathbf{x}_i^\mu/dt = \partial H_{I_0}/\partial \dot{p}_\mu^i - v^\sigma A_\sigma^\mu(\mathbf{x}_i), \quad d\alpha^\sigma/dt = v^\rho G_\rho^\sigma \quad (4.10)$$

contain the indefinite velocities v^ρ .

Between the special H_{I_0} and L_I (4.6) there is a many-valued correspondence. The most conventional way to construct an H_{I_0} is as follows: First, we can perform the Legendre transformation to the canonical system keeping the variables in (P). Uniquely the (P)-Hamiltonian $H_{0p}(\bar{\mathbf{x}}, \bar{\mathbf{p}})$ can be obtained. It is then transformed into (I) by a canonical transformation consistent with the point transformation (4.1). The latter Hamiltonian is identified with H_{I_0} . In this course

$$W = \sum S^\mu(\mathbf{x}_i; \alpha) \bar{p}_\mu + \alpha^\sigma \bar{\beta}_\sigma \quad (4.11)$$

is the generating function of the transformation.

The inverse transformation generated by the same function (4.11) brings the (I)-Hamiltonian up to the terms $v^\rho \phi_\rho$ back into the (P)-Hamiltonian

$$H_p \equiv H_{0p} + v^\sigma G_\sigma^\rho(\bar{u}) \bar{\beta}_\rho \quad (4.12)$$

The equations of motion from (4.12)

$$d\bar{u}^\sigma/dt = v^\rho G_\rho^\sigma(\bar{u}), \quad d\bar{\beta}_\sigma/dt = -(v \partial G/\partial \alpha \beta)_\sigma, \quad (4.13)$$

with subsidiary conditions $\phi_\sigma = G_\sigma^\rho(\bar{u}) \bar{\beta}_\rho = 0$, are obtained with respect to \bar{u} and $\bar{\beta}$ from (4.12). Because of $\det|G(\alpha)| \neq 0$, the condition is equivalent to $\bar{\beta} = 0$, and consequently $d\bar{\beta}/dt = 0$. None of the variables v are there in the equations for $\bar{\mathbf{x}}$'s and $\bar{\mathbf{p}}$'s. Since no information about v 's is thus obtained from the equations, the constraint $d\bar{u}^\sigma/dt = v^\sigma = 0$ does not alter them. And we have no longer to be interested in the world of the variables \bar{u} and $\bar{\beta}$. Consequently we have arrived at the original representation with none of such conditions. This special choice of v 's can be applied also to the variables in (I). One finds a complete analogy in the conventional choice of Fermi or Lorentz gauge in the problem of electrodynamics.

The conservation law

$$d\bar{\phi}_\sigma/dt = 0 \quad (4.14)$$

is easily verified in (P). Correspondingly

$$d\phi_\sigma/dt = 0 \quad (4.14)'$$

holds also in (I), for (4.11) is independent of time. The quantity in (I),

$$\Omega \equiv \phi_\sigma \xi^\sigma \quad (4.15)$$

operates as the generating function of infinitesimal gauge transformation (4.4). So as to guarantee the group theoretical properties ϕ_σ must satisfy, therefore,

$$C_{\sigma\rho}^\tau \phi_\tau = [\phi_\sigma, \phi_\rho]; \quad (4.16)$$

that is to say, $[\phi_\sigma, \phi_\rho] = 0$ is not any new condition other than $\phi_\sigma = 0$. The quantity C is usually called the structure constant of the group; it contains neither x 's nor α 's, but it is a set of pure constant numbers proper to the group. (See appendix I.)

§ 5. Interrelation to the hydrodynamics

In this paragraph the interaction potential term V is dropped in order to avoid inessential complexity. The general form of (I) Hamiltonian is

$$H_I \equiv (1/2M) \sum g^{\mu\nu}(\mathbf{x}_i; \alpha) p_\mu^i p_\nu^i - v^\sigma (\sum A_\sigma^\mu(\mathbf{x}_i) p_\mu^i - G_\sigma^\rho(\alpha) \beta_\rho), \quad (5.1)$$

where

$$g_{\lambda\mu}(\mathbf{x}_i; \alpha) g^{\kappa\nu}(\mathbf{x}_i; \alpha) = \delta_{\lambda\mu}^\nu, \quad g_{\mu\nu} = \frac{\partial S_\kappa}{\partial x^\mu} \cdot \frac{\partial S^\kappa}{\partial x^\nu}. \quad (5.2)$$

(Notice \mathbf{x} -space is no longer Euclidian!) From (5.1) the equations of motion

$$dx_i^\mu/dt = (1/M) g^{\mu\nu}(\mathbf{x}_i, \alpha) p_\nu^i - v^\sigma A_\sigma^\mu(\mathbf{x}_i) \quad (5.3)$$

are obtained. The linear combination of (5.3)

$$\sum_i g_{\mu\nu}(\mathbf{x}_i; \alpha) A_\rho^\mu(\mathbf{x}_i) dx_i^\nu/dt = (1/M) \sum A_\rho^\mu(\mathbf{x}_i) p_\mu^i - v^\sigma \Gamma'_{\sigma\rho},$$

where

$$\Gamma'_{\sigma\rho} \equiv \sum_i g_{\mu\nu}(\mathbf{x}_i, \alpha) A_\sigma^\mu(\mathbf{x}_i) A_\rho^\nu(\mathbf{x}_i)$$

becomes equal to

$$(1/M) G_\rho^\sigma(\alpha) \beta_\sigma - v^\sigma \Gamma'_{\sigma\rho}, \quad (5.4)$$

owing to the subsidiary condition. It turns, by means of a special choice of v such as

$$v^\sigma = \Gamma'^{\sigma\rho} (G_{\rho i}^{\tau\tau} M - \Gamma_\rho'),$$

where $\Gamma'^{\sigma\tau} \Gamma'_{\tau\rho} = \delta_{\rho}^\sigma$, into

$$v_\rho' = \sum g_{\mu\nu}(\mathbf{x}_i, \alpha) A_\rho^\nu(\mathbf{x}_i) dx_i^\mu/dt.$$

It is the sum of parallel component of velocities to that of the organized motion described by α 's. Hitherto, the variables α^σ have been regarded as dynamical variables, but, giving up Lagrange picture, we may take Euler picture. α 's are then regarded as merely parameters. By the stand-point (5.4) can be re-written as

$$v_\rho' = \sum A_\rho^\mu(\bar{\mathbf{x}}_i) d\bar{x}_i^\mu/dt. \quad (5.4)'$$

Consequently the quantity A_ρ^μ has the physical meaning of a certain mode of velocity field. The theory of Pfaff's problem shows that $A_\rho^\mu(\bar{\mathbf{x}})$ can be, in general, decomposed into

$$A_p^\mu(\bar{\mathbf{x}}) = \partial \phi_p / \partial \bar{x}_\mu + A_p(\mathbf{x}) \partial M / \partial \bar{x}_\mu. \quad (5.5)$$

In addition (5.4) can be expressed in the hydrodynamical manners as

$$v_p' = \int A_p^\mu(\bar{\mathbf{x}}) d\bar{x}_\mu / dt \cdot \rho(\bar{\mathbf{x}}) d\bar{v}, \quad (5.6)$$

$$\rho(\bar{\mathbf{x}}) = \sum_i \delta(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i). \quad (5.7)$$

The right-hand side of (5.4) can be integrated in some cases where $A_p = 0$, but cannot be in other cases where $A_p \neq 0$. We call the one 'vortex case', and the other 'irrotational' case; either of our preceding the examples falls into the latter category. In the vortex case a sort of vortex motion will be observed on the flow of the many-particle system. In general A_p and M are not uniquely defined, but they may be restricted by a condition such as $\nabla A_p \nabla M = 0$.

§ 6. Collective representation (Principle of stiffening)

The first representation will not be suitable. There we have r subsidiary conditions including both p 's and β 's. The collective description will, however, be important when we are much more interested in the behaviours of α 's and β 's than those of x 's and p 's. It is desirable to ban out β 's from ϕ 's by a suitable canonical transformation.* Let the generating function of the transformation put

$$W = \sum_i x_i^\mu P_\mu^i + A^\sigma(\mathbf{x}, \alpha) B^\sigma, \quad (6.1)$$

where (X, P, A, B) are the new variables.

First, $A^\sigma(\mathbf{x}, \alpha)$'s have to satisfy the differential equations

$$\sum_i A_p^\mu(\mathbf{x}_i) \partial A^\sigma / \partial x_i^\mu - G_p^\sigma(\alpha) \partial A^\sigma / \partial \alpha^\tau = 0, \quad (6.2)$$

so that ϕ_p may turn into the form free from B 's. The conditions (6.2) are, however, too weak to determine the concrete functional forms of A 's because their only requirement is that A 's should never depend on \mathbf{x} 's and α 's but only through $\bar{\mathbf{x}}$'s.

The second restriction for A 's is that the transformation $\alpha \rightarrow A$ should be unique and reversible; namely,

$$\det |\partial A^\sigma / \partial \alpha^\rho| \neq 0. \quad (6.3)$$

We can solve $A = A(\mathbf{x}, \alpha)$ with respect to α 's once the $A(\mathbf{x}, \alpha)$'s are set up, and $\alpha = \alpha(\mathbf{x}, A)$ satisfies

$$\sum_i A_p^\mu(\mathbf{x}_i) \partial A^\sigma / \partial x_i^\mu + G_p^\sigma(\alpha) = 0. \quad (6.2)'$$

the new variable are expressed by the old ones through the expressions:

$$\begin{aligned} x_i^\mu &= X_i^\mu, & p_\mu^i &= P_\mu^i + (\partial A^\sigma / \partial x_i^\mu) B_\sigma, \\ A^\sigma &= A^\sigma(\mathbf{x}, \alpha), & \beta_\sigma &= (\partial A^\rho / \partial \alpha^\sigma) B_\rho. \end{aligned} \quad (6.4)$$

* Such a procedure is in the same manner as in the works of other authors. Cf. ref. (11), (12).

Also the Hamiltonian (5.11) is transformed into

$$H = (1/2M) \sum g^{\mu\nu}(\mathbf{x}; \alpha) P_\mu^i P_\nu^i + (1/2M) \sum g^{\mu\nu}(\mathbf{x}; \alpha) \left(\frac{\partial A^\sigma}{\partial x_i^\mu} \cdot \frac{\partial A^\rho}{\partial x_j^\nu} \right) B_\sigma B_\rho \\ + (1/M) \sum (g^{\mu\nu}(\mathbf{x}; \alpha) (\partial A^\sigma / \partial x_i^\mu) P_\nu^i) B_\sigma + \tilde{V}(\mathbf{X}, \mathbf{A}) - v^p \tilde{\phi}_p. \quad (6.6)$$

It is understood that the old variables are all expressed in terms of the new ones.

In the third place it is expected the functional form of A 's to be so constructed as to enable the third term of (6.6) to amalgamate the fourth. The partial derivatives of A 's are written in the form

$$\partial A^\sigma / \partial x_i^\mu = \kappa \{ \gamma^{\sigma\rho} g_{\mu\nu}(\mathbf{x}; \alpha) A_\rho^\nu(\mathbf{x}_i) \} + A_\mu^\sigma(\mathbf{x}_i)' \quad (6.7)$$

where γ is the ground metric tensor of α -space, κ is a suitable proportional factor, and A_μ^σ is a remainder that cannot be expressed in an analogous form. At $\alpha=0$ $\gamma^{\sigma\rho}$ is equal to $\delta^{\sigma\rho}$. In solving (6.7) as regards A^σ we may operate at the origin of α -space where $\mathbf{X}=\mathbf{x}=\bar{\mathbf{x}}$, because there appear none of α -derivatives, and because A 's are dependent only on \mathbf{x} 's. Thus we obtain

$$\partial A_\sigma / \partial \bar{x}_i^\mu = \kappa A_{\sigma\mu}(\bar{\mathbf{x}}_i) + A_{\sigma\mu}(\bar{\mathbf{x}}_i)'. \quad (6.7)'$$

Here we may relax our exertions to the positions of indices.

On more somewhat categorical statement on $A(\mathbf{x}_i, \alpha)$'s is that they are assumed to be symmetric functions with respect to the co-ordinates of particles.

According to the discussion appeared in § 5 there hold the identities

$$A_{\sigma\mu}(\bar{\mathbf{x}}) \equiv \partial \phi_\sigma / \partial \bar{x}_\mu + A_\sigma(\bar{\mathbf{x}}) \partial M / \partial \bar{x}_\mu.$$

So we obtain

$$A_\sigma = \sum \kappa \{ \phi_\sigma(\bar{\mathbf{x}}_i) + A_\sigma(\bar{\mathbf{x}}_i) M(\bar{\mathbf{x}}_i) \}. \quad (6.8)$$

$A'_{\sigma\mu}$ is identified with

$$A'_{\sigma\mu} \equiv \kappa (\partial A_\sigma / \partial \bar{x}_{\mu i}) M(\bar{\mathbf{x}}_i). \quad (6.9)$$

Then the third term of (6.6) becomes

$$\sum (\kappa/M) \gamma^{\sigma\rho} \{ A_\rho^\mu(\mathbf{X}_i) P_\mu^i + g^{\mu\nu}(\mathbf{X}_i, \alpha) (\partial A_\rho / \partial x_i^\mu) M \cdot P_\nu^i \} B_\sigma. \quad (6.10)$$

The former term of (6.10) amalgamates into \tilde{v}_σ , and the latter, together with such a kind of terms arising from $(1/2M) \sum (\partial A^\sigma / \partial x_i^\mu) g_{\mu\nu} (\partial A^\rho / \partial x_j^\nu) B_\sigma B_\rho$ of (6.6), transmutes P 's in its first term to

$$P_\mu^{i*} \equiv P_\mu^i + \kappa \gamma^{\sigma\rho} g_{\lambda\mu}(\mathbf{X}_i, \alpha) (\partial A_\rho / \partial x_i^\lambda) B_\sigma \cdot M(\mathbf{X}_i, \alpha), \quad (6.11)$$

and consequently the transformed Hamiltonian $H_{II} = H_{II0} - v'^\sigma \tilde{\phi}_\sigma$

$$H_{II0} \equiv (1/2M) \sum (P_i^* \cdot g \cdot P_i^*) + (\kappa^2/2M) (\mathbf{B} \cdot \mathbf{I}' \cdot \mathbf{B}) + \tilde{V}(\mathbf{X}, \mathbf{A})$$

$$\text{with subsidiary conditions; } \tilde{\phi}_\sigma \equiv \sum A_\sigma^\mu(\mathbf{X}_i) P_\mu^i = 0 \quad (6.12)$$

is obtained, where $v'^\sigma = v'^\sigma - (\kappa/M) (\gamma \cdot \mathbf{B})^\sigma$. We shall call this representation "The

second representation" or (II).

Now, again the principle of stiffening will be applied. The subsidiary conditions are replaced by a set of constraints. We see

$$d\tilde{\chi}_\sigma \equiv \sum g_{\mu\nu}(\mathbf{X}_i, \alpha) A_\sigma^\mu(\mathbf{X}_i) dX_i^\nu = -v'^\rho I'_{\rho\sigma} dt \quad (6.13)$$

as a consequence of $\tilde{\phi}_\sigma = 0$. (6.13) defines the affine co-ordinates $\tilde{\chi}_\sigma$'s. Owing to the arbitrariness of v' 's we may put them to vanish, and then (6.13) turns into

$$d\tilde{\chi}_\sigma = 0. \quad (6.13)'$$

Next, let us consider another dynamical system with the same Hamiltonian as (6.12) but stiffened by the constraint (6.13)'. $\phi_\sigma = 0$ becomes then a consequence of the constraint; namely

$$d\tilde{\chi}_\sigma/dt \equiv \sum g_{\mu\nu}(\mathbf{X}_i, \alpha) A_\sigma^\mu(\mathbf{X}_i) dX_i^\nu/dt = \phi_\sigma/M = 0. \quad (6.14)$$

Tentatively add the terms of passive forces, whether holonomic or not, to the equations of motion with certain measures of those forces u^σ , and

$$d\tilde{\phi}_\sigma/dt = [H_{II0}, \tilde{\phi}_\sigma] - u^\rho [\tilde{\chi}_\rho, \phi_\sigma] = u^\rho I'_{\sigma\rho} \quad (6.15)$$

are obtained. (6.15) and (6.14) show that u^σ should be equal to naught. Every term of these passive forces vanishes. Our new dynamical system is, even in general, equivalent to the original one.

By putting the generating function to be

$$W = \sum \kappa \gamma^{\sigma\rho} (\phi_\rho(\alpha) + A_\rho(\mathbf{x}_i, \alpha) M(\alpha)) B_\sigma + \sum x_i^\mu P_\mu^i, \quad (6.16)$$

we try the third transformation $(\mathbf{X}, \mathbf{A}) \rightarrow (\mathbf{x}, \alpha)$. (6.16) is apparently in the same form as (6.1). The \mathbf{x} 's in ϕ and M are, however, regarded as parameters by taking account of (6.13)'. This final representation has a manageable form of Hamiltonian

$$H_c \equiv (1/2M) \sum g^{\mu\nu}(\mathbf{x}, \alpha) p_\mu^i p_\nu^i + (1/2M) \Gamma^{\rho\sigma} L_\rho L_\sigma + V(\mathbf{x}, \alpha), \quad (6.17)$$

subject to $d\chi_\sigma \equiv \sum g_{\mu\nu}(\mathbf{x}_i, \alpha) A_\sigma^\mu(\mathbf{x}_i) dx_i^\nu = 0$. The quantity L_σ defined by

$$L_\rho \equiv G_\rho^\sigma(\alpha) \beta_\sigma \quad (6.18)$$

satisfies the algebraic relation as follows;

$$[L_\rho, L_\sigma] = C_{\rho\sigma}^\tau L_\tau. \quad (6.19)$$

This "Collective representation", or (C), is not only very much handy to pull out the collective mode of motion explicitly, but also is obtained without making such a detour $\Psi(\mathbf{P}) \rightarrow (\text{I}) \rightarrow (\text{II}) \rightarrow (\text{C})$: i.e., we can at once get H_c from H_p if any transformation is imagined corresponding to certain mode of motion. The second term of (6.17) can be considered as the collective kinetic energy. It originates from the fact that the dynamical properties were explicitly transferred into the transformation parameters from the variables of individual particles.

§ 7. Digression on approximate treatment

Hitherto we have not much concerned ourselves with the question of physical correspondence, but only with the mathematical introduction of variables to describe certain modes of collective motion by suitable continuous transformation. All of such introduced variables, however, do not necessarily possess the physical significance. They would be conditioned by some additional statements related to the approximation procedures so as to be physical notions. Fundamentally the collective motion as a physical conception may be anticipated in the following cases: (i) There may happen to be sufficient numbers of integrals of motion containing only the variables α 's and β 's. Especially it will be the most favourable case if the (C)-Hamiltonian is divided into two terms as $H_1(x, p) + H_2(\alpha, \beta)$. Then the collective motion can be treated completely independently of individual motion. (ii) Besides such an ideal case even interferences between these two motions may be allowed so far as they are sufficiently small. In such a case we can often appeal to the so-called adiabatic approximations.

In general, H_c consists of three terms; the individual kinetic energy T_p , the collective kinetic energy T_c , and the interaction potential $V(x, \alpha)$. Such a clear-cut classification is, however, not always expected. Each of two kinetic energy terms contains the co-ordinates of another through y or I' , while the momenta are separated. (y & I' are pure numbers only in the particular case of c.m.). Such an ideal case as (i) seldom occurs.

For example, from the Hamiltonian (3.10) the equation of motion

$$dI''/dt = (2/M) \left[\left\{ \sum_i (x_i p'_x) + \sum_j (y_j p'_y) \right\} + (\alpha_0 \beta) \right] \quad (7.1)$$

is obtained for I'' appeared in § 3. (Owing to the constraint $\sum_i (x_i p'_x) = \sum_j (y_j p'_y)$ holds, but it hardly simplifies the matter.) The term bundled by the curly bracket is a sum of multitude of inner products $(x_i \cdot p_i)$. If the system has an isotropic distribution in the phase space, then (7.1) becomes effectively equivalent to

$$dI'_c/dt = (2/M) (\alpha_0 \beta). \quad (7.2)$$

Thus we may replace I'' by another function I'_c independent of x and y but α . We may take I'_c equal to $I''(x_0, \alpha)$ where x_0 's are certain time-average values of x 's. Thus T_c turns into a separated form

$$T_c \equiv (1/2M) (1/I'_c) (\alpha_0 \beta)^2. \quad (7.3)$$

Similarly the genuine $I''_{\sigma\rho}$ can often be effectively replaced by $I'_{c\sigma\rho}$ independent of x 's under an analogous condition. It seems, however, too severe. Some advanced considerations will be necessary to proceed further.

The quantity $I'_{\sigma\rho}$ is a natural generalization of the tensor of inertia, appeared also in the Tomonaga's work. $I'_{\sigma\rho}$ is almost proportional to $1/N$. To the zeroth approximation T_c is dropped when an adiabatic approximation can be applied. Just as in the theory of molecules first we solve the equations of motion by a reduced Hamiltonian H

$$H_R \equiv T_P + V \quad (7.4)$$

under suitable initial conditions, and then an energy integral $U(\alpha)$ is obtained at a certain point of α -space. $U(\alpha)$ is divided into two parts as

$$U(\alpha) = U_i(\alpha) + U_c(\alpha). \quad (7.5)$$

$U_i(\alpha)$ is variable with the initial conditions, while $U_c(\alpha)$ is independent of them. The ideal case will occur if $\partial U_i / \partial \alpha = 0$ holds everywhere in the α -space, and then the collective motion induced by α is well-defined. The three criteria for V at the end of § 2 fall into such a case. Otherwise, only in the neighbourhood of the extrema of U_i it will be well-defined. So as to respond to a stable collective motion these extrema must be minimum, the minimum points of U_c must be sufficiently near these points, and the curvature of U_i at these points must be flat compared to that of U_c . To avoid the ambiguities concerning the determination of U_c requires to consider some statistical concepts such as the equilibrium state of the system and so on.

As an example let us treat the case of § 3 assuming the forms of V such as

$$V(\bar{\mathbf{x}}_1 \cdots \bar{\mathbf{x}}_N) = \sum_{i>j} V(|\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j|), \quad (7.6)$$

$$V(r) = D \{1 - \exp \kappa(r_e - r)\}^2, \quad (7.7)$$

where D , κ and r_e are the adjustable parameters. V is expanded in the power series about $\alpha=0$

$$V(r) = V(r) + \alpha r \cos \phi V'(r) - (\alpha^2/8) \{3rV'' - r^2V'''\} + O(\alpha^2), \quad (7.8)$$

after the transformation as (4.1)'. Since the second and the higher order term $O(\alpha)^2$ are proportional to $\cos n\phi$, and since ϕ_{ij} is the polar angle of the interparticle radius vector \mathbf{r}_{ij} , the sum vanishes most probably, and consequently

$$\sum V(r) = \sum V(r) - (\alpha^2/8) \sum \{3rV'' - r^2V'''\} \quad (7.9)$$

holds as a good approximation. Further provided that every particle is in the vicinity of the equilibrium point; i.e., the system is in the lower energy equilibrium state, then we have in consequence of short range nature of forces

$$(\alpha^2/8) \langle \sum (3V''r - r^2V''') \rangle_{AV} \approx (\alpha^2/8) N r_e^2 V''(r_e). \quad (7.10)$$

It corresponds to U_c , and here U_i is negligible. The frequency of random small oscillation is of the order of $f_p \approx \kappa \sqrt{D}$, while that of organized oscillation is of the order of $f_c \approx r_e f_p / R_0 \sqrt{N}$, where R_0 is the average linear dimension of the system, so that the adiabatic approximation is good. Thus the Morse-type two-body potential gives rise to the well-defined incompressible motion. Also it is easy to prove that a δ -type potential gives rise to the dilatation as well as the incompressible motion.

§ 8. Two more examples

(1) Two dimensional rotator :

By virtue of the general theory so far advanced the formulation is now systematic

enough to be applicable even to more complicated cases. In (3.1), by putting $\alpha_1 = \alpha_3 = 0$, we obtain the plane rotation

$$\bar{x} = \alpha_0 x - \alpha_0 y, \quad \bar{y} = \alpha_0 x + \alpha_0 y, \quad \alpha_0^2 + \alpha^2 = 1. \quad (8.1)$$

Since the original (P)-Hamiltonian is rotationally invariant, the (I)-Hamiltonian H_{I_0} has the same form as H_p

$$H_{I_0} \equiv (1/2M) \sum (p_x^2 + p_y^2) + V(\mathbf{x}), \quad (8.2)$$

but with a subsidiary condition

$$\phi \equiv \sum_i (x_i p_y^i - y_i p_x^i) - \alpha_0 \beta = 0. \quad (8.3)$$

The related constraint in (C) is of course non-integrable. It falls into the vortex case. The (C)-Hamiltonian is written down as

$$H_c \equiv (1/2M) \sum (p_x^2 + p_y^2) + V(\mathbf{x}) + (1/2M I') (\alpha_0 \beta)^2, \quad (8.4)$$

where I' , the moment of inertia, is given by

$$I' \equiv \sum (x_i^2 + y_i^2). \quad (8.5)$$

The third term of (8.4) represents the kinetic energy of a rigid rotator in two dimensions if the constraints

$$d\chi \equiv \sum (x_i dy_i - y_i dx_i) = 0 \quad (8.6)$$

in its general sense is replaced by the more stringent conditions $dx_i = dy_i = 0$, so that I' is exactly regarded as a constant number.

(2) *Three-parametric incompressible motion:*

A little more interesting example is three-parametric incompressible motion in two dimensions proposed already in § 3. The two-dimensional unimodular group constitutes a non-commutative group; a point that invites a little complication. We can, however, immediately write down the (C)-Hamiltonian without making a detour to the passing-through stages. The (C)-Hamiltonian

$$H_c \equiv (1/2M) \sum g^{\mu\nu}(\alpha) p_\mu^i p_\nu^i + V(\mathbf{x}, \alpha) + (1/2M) I'^{\alpha\beta} L_\alpha L_\beta \quad (8.7)$$

is then obtained with three constraints

$$\begin{aligned} d\chi_1 &\equiv g_{21} \sum (x_i dy_i + y_i dx_i) = 0, \\ d\chi_2 &\equiv g_{21} \sum (x_i dy_i - y_i dx_i) = 0, \\ d\chi_3 &\equiv \sum (g_{11} x_i dx_i - g_{22} y_i dy_i) = 0, \end{aligned} \quad (8.8)$$

where

$$\begin{aligned} g_{11} &= g^{22} = 1 + 2(\alpha_0 \alpha_3 + \alpha_1 \alpha_2) + 2(\alpha_1^2 + \alpha_3^2), \\ g_{22} &= g^{11} = 1 - 2(\alpha_0 \alpha_3 + \alpha_1 \alpha_2) + 2(\alpha_1^2 + \alpha_3^2), \\ g_{12} &= g_{21} = -g^{21} = -g^{12} = 2(\alpha_0 \alpha_1 - \alpha_2 \alpha_3), \end{aligned} \quad (8.9)$$

and

$$\alpha_0^2 - \alpha_1^2 + \alpha_2^2 - \alpha_3^2 = 1. \quad (8 \cdot 10)$$

Practically we may drop the higher power than the third order of α_1 and α_3 , and in the neighbourhood of $\alpha_1 = \alpha_3 = 0$ I^2 's are replaced by

$$I^{11} = I^{22} = I^{33} = 1/NR_0^2, \quad I^{ik} = 0 (i \neq k); \quad NR_0^2 \cong \sum (x_i^2 + y_i^2) \quad (8 \cdot 11)$$

for a large value of N . And approximately

$$T_c^0 \equiv (1/2NR_0^2M) (L_1^2 + L_2^2 + L_3^2) \quad (8 \cdot 12)$$

is used in place of the collective kinetic energy T_c . The structure constant C is

$$C_{12}^3 = C_{23}^1 = C_{13}^2 = -C_{21}^3 = -C_{32}^1 = -C_{31}^2 = 2, \quad \text{other elements} = 0. \quad (8 \cdot 13)$$

Consequently the algebraic relations for L 's

$$[L_1, L_2] = 2L_3, \quad [L_2, L_3] = 2L_1, \quad [L_3, L_1] = -2L_2 \quad (8 \cdot 14)$$

are deduced, where

$$\begin{aligned} L_1 &\equiv \alpha_0\beta_1 - (\alpha_2\beta_3 - \alpha_3\beta_2), \\ L_2 &\equiv \alpha_0\beta_2 + (\alpha_3\beta_1 - \alpha_1\beta_3) = L_2^R + L_2^B, \\ L_3 &\equiv \alpha_0\beta_3 - (\alpha_1\beta_2 - \alpha_2\beta_1). \end{aligned} \quad (8 \cdot 15)$$

The first term of L_2 clearly means the momentum of rigid rotation, and the second is signified as the momentum of Bohr rotation. In the neighbourhood of $\alpha_1 = \alpha_3 = 0$ $L_1^2 + L_2^2 = \beta_1^2 + \beta_2^2$ means the kinetic energy of deformation oscillation. It contains such a term as L_2^{R2}/α^2 , ($\alpha^2 = \alpha_1^2 + \alpha_3^2$), which is the main term of Bohr rotation. We can expand also potential energy in the power series of α_1 and α_3 as

$$V \approx \sum V(r_{ij}) + (N/8)\alpha^2 \cdot r_e^2 V''(r_e), \quad (8 \cdot 16)$$

assuming the Morse-type. The Hamiltonian is then circular symmetric in the $\alpha_1 - \alpha_3$ plane, and the quantity L_2^R becomes a constant of motion labelling the energy integral in a certain manner as well as another constant of motion L_2^B ; the lower order of H_c contains α_0 nowhere. The ordinary space rotation corresponds to the two-valued representation of Bohr rotation. In words of the quantum theory L_2^R has eigenvalues of even numbers $0, \pm 2, \pm 4, \dots$, while L_2^B has those of $0, \pm 1, \pm 2, \dots$. The difference of the moment of inertia $\alpha^2 I''$ and I'' concerning each of two terms of angular momenta effects the level density of energy. This classification will become much more clear-cut if N is large.

§ 9. Conclusions

We have developed the general theory of collective motion by using the auxiliary variables. The properties of collective motions are attributed to the properties of transformations; that is, the auxiliary variables to describe a certain collective mode of motion first

enter as transformation parameters. Consequently several subsidiary conditions have been introduced systematically. These are dealt along the group theoretical way. The conditions are replaced by the same numbers of Pfaffian constraints in (C), the final representation. If any of these Pfaffians are expressed in the integrated forms, our theory is equivalent to the theory of Miyazima and other authors. The Pfaff's form is, however, not always integrable. Our consideration contains such a more general case with non-integrable constraints, viz. 'vortex cases'.

We have only to find the quantities g , I , A and C in order to write down the (C)-Hamiltonian from the given (P)-Hamiltonian if a suitable transformation is imagined concerning a mode of collective motion. It is a point that the theory simplifies the treatments very much. On the other hand the unavoidable constraints make complexity apparently. They are, however, not so much important when N is large compared with r , practically. In order to have a long life-time of the proposed collective motion the transformation property of Hamiltonian plays an important part: for example, the collective motion, as introduced, will have a good physical significance if the transformed Hamiltonian can be treated by means of the adiabatic approximation. The goodness of the approximation brings the physical significance of the collective motion.

Such a treatment may be applied also to the field theory. Some additional fields will be derived by a certain generalized gauge transformation corresponding to the transformation (4.1). The excessive fields, appeared in the works of Yang and Mills and of Utiyama, will get thus a certain physical meaning.²¹⁾²²⁾

The author should like to express his thanks to Professors K. Husimi, S. Tomonaga, T. Miyazima, Drs. T. Nisiyama and T. Marumori, for their valuable discussions; — particularly to Prof. Husimi for his detailed advices and encouragements through about half a year for revision. In addition he also should like to manifest many thanks to the authorities of the Research Institute for Fundamental Physics, Kyoto University who had given us much convenience for this work.

Appendix I.

Group theoretical properties of continuous transformation

Let the set of transformations that removes \bar{x}^μ to x^μ by

$$T_\alpha : \quad x^\mu = f^\mu(\bar{x}, \alpha) \quad (\text{A} \cdot 1)$$

be an r -parameter (α^n) continuous group defined over x -space. We can obtain the differential equations that f^μ would satisfy as functions of α^n . Without spoiling generality,

$$T_0 = 1 : \quad \bar{x}^\mu = f^\mu(\bar{x}, 0) \quad (\text{A} \cdot 2)$$

can be assumed. Put the inverse transformation of T_α

$$T_\alpha^{-1} : \quad \bar{x}^\mu = f^\mu(x, \phi(\alpha)), \quad (\text{A} \cdot 3)$$

where $\phi^o(\alpha)$ tends to $-\alpha^o$ with $\alpha^o \rightarrow 0$, and define the composition function $\varphi^o(\alpha, \alpha')$ by

$$\mathbf{T}_\alpha \cdot \mathbf{T}_{\alpha'} : \quad x^\mu = f^\mu(f(\bar{x}, \alpha), \alpha') = f^\mu(\bar{x}, \varphi). \quad (\text{A} \cdot 4)$$

Corresponding to the operation of differentiation in ordinary sense, the differential of transformation is defined

$$\Delta \mathbf{T}_\alpha = \mathbf{T}_{\alpha + \Delta \alpha} \mathbf{T}_\alpha^{-1} : \quad x^{\mu'} = f^\mu(f(x, \psi(\alpha)), \alpha + \Delta \alpha), \quad (\text{A} \cdot 5)$$

where $x^{\mu'} = f^\mu(\bar{x}, \alpha + \Delta \alpha)$. According to $\varphi^\sigma(\alpha, \psi(\alpha)) = 0$, and $\lim_{\Delta \alpha \rightarrow 0} \varphi^\sigma(\psi(\alpha), \alpha + \Delta \alpha) = 0$, we get

$$\xi^\sigma = A_p^\sigma(\alpha) \Delta \alpha^p, \quad (\xi^\sigma; \text{infinitesimal})$$

where

$$A_p^\sigma = \partial \varphi^\sigma(\alpha, \alpha') / \partial \alpha'^p |_{\alpha' \rightarrow \alpha}, \quad (\text{A} \cdot 6)$$

by putting $\Delta \varphi^\sigma = \xi^\sigma$. Clearly $x^\mu = x'^\mu$ when $\Delta \alpha = 0$. $\Delta \alpha$ being infinitesimal, $\Delta x^\mu = x'^\mu - x^\mu$ is also infinitesimal. Put

$$A_p^\mu(x) \equiv \partial f^\mu / \partial \alpha^p |_{\alpha \rightarrow 0}, \quad (\text{A} \cdot 7)$$

and we have

$$dx^\mu = A_o^\mu(x) \xi^\sigma = A_p^\mu(x) A_o^p(\alpha) d\alpha^\sigma, \quad (\text{A} \cdot 8)$$

or, in a familiar form of Lie's fundamental equation,

$$\partial f^\mu / \partial \alpha^\sigma = A_p^\mu(x) A_o^p(\alpha). \quad (\text{L} \cdot \text{I})$$

By using the inverse expression

$$\mathbf{T}_\alpha^{-1} : \quad \bar{x}^\mu = S^\mu(x, \alpha) = f^\mu(x, \psi), \quad (\text{A} \cdot 3)'$$

we get instead of (L·I)

$$A_o^\sigma(\alpha) A_o^\mu(x) \partial S^\nu / \partial x^\mu - \partial S^\nu / \partial \alpha^\sigma = 0. \quad (\text{L} \cdot \text{I})'$$

It is worth noticing that the right-hand side of (L·I) takes such a special form as the sum of products of two factors; one of them is the function of x 's only, and another is of α 's only. As is easily seen $A(0) = \mathbf{I}$, and $\det|A(0)| = 1$, so $\det|A|$ has a certain non-vanishing domain owing to the continuity. Consequently α has a certain domain of variability. Further the composition equation

$$\alpha^{\sigma''} = \varphi^\sigma(\alpha, \alpha') \quad (\text{A} \cdot 9)$$

constructs a continuous group, too. By replacing \bar{x} , α and x by α , α' and α'' in the above analysis, we see φ^σ satisfies the Lie's equation

$$\partial \varphi^\sigma / \partial \alpha'^p = G_p^\sigma(\alpha) A_p^\sigma(\alpha'), \quad (\text{L} \cdot \text{I})''$$

where G is defined in the same manner as A . It is clear that

$$G_p^\sigma(\alpha) A_o^\sigma(\alpha) = \partial_o^\sigma;$$

namely

$$\mathbf{G} = \mathbf{A}^{-1}. \quad (\text{A} \cdot 10)$$

It results (4·2).

According to the complete integrability of f^μ

$$\partial^2 f^\mu / \partial \alpha^\sigma \partial \alpha^\rho = \partial^2 f^\mu / \partial \alpha^\rho \partial \alpha^\sigma,$$

we can deduce

$$A_\rho^\mu \partial A_\sigma^\nu / \partial x^\mu - A_\sigma^\mu \partial A_\rho^\nu / \partial x^\mu = C_{\rho\sigma}^\tau A_\tau^\nu, \quad (\text{A} \cdot 11)$$

where C is defined by

$$C_{\rho\sigma}^\tau \equiv G_\rho^\omega G_\sigma^\nu (\partial A_\omega^\tau / \partial \alpha^\nu - \partial A_\nu^\tau / \partial \alpha^\omega). \quad (\text{A} \cdot 11)'$$

It is easily proved that the function C is a set of pure constant numbers, by virtue of (A·11) and (A·11)'. It is called the structure constant of the group. The differential operators defined by

$$\lambda_\sigma \equiv A_\sigma^\mu \partial / \partial x^\mu, \quad (\text{A} \cdot 12)$$

satisfy the commutation relation according to (A·11)

$$[\lambda_\sigma, \lambda_\rho] = C_{\sigma\rho}^\tau \lambda_\tau. \quad (\text{A} \cdot 13)$$

Also we have from (A·11)'

$$[L_\sigma, L_\rho] = C_{\sigma\rho}^\tau L_\tau, \quad (\text{A} \cdot 13)'$$

where

$$L_\sigma \equiv G_\sigma^\rho(\alpha) \partial / \partial \alpha^\rho. \quad (\text{A} \cdot 12)'$$

Consequently we get

$$[\phi_\sigma, \phi_\rho] = C_{\sigma\rho}^\tau \phi_\tau, \quad (\text{L} \cdot \text{II})$$

where

$$\phi_\sigma = \lambda_\sigma - L_\sigma.$$

The second theorem of Lie (L·II) has the same content as (16).

Appendix II.

On Dirac's generalized Poisson bracket (G.P.B.)

In the final representation (C) we have r differential forms as constraint for the variables of individual particles, x_i^μ . They are conventionally taken into account by modifying the equations of motion by means of additional passive forces. They will, however, be introduced logically by altering the meanings of differentiation operation. This concept makes easier to quantize the internal system. The idea lies in modifying the Poisson bracket after the manner of P.A.M. Dirac.²³⁾ In the original sense, P.b.'s are subject to

$$\begin{aligned} \text{(i)} \quad & [\xi, \eta] = -[\eta, \xi], \\ \text{(ii)} \quad & [\xi, \eta + \zeta] = [\xi, \eta] + [\xi, \zeta], \\ \text{(iii)} \quad & [\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = 0, \end{aligned} \quad (\text{A} \cdot 14)$$

for any dynamical variables ξ , η and ζ . The notion P.b.'s may be extended to include quantities depending on x 's and α 's, of which only the differentials are defined, but which are expressible in terms of neither x 's nor α 's. We postulate these generalized P.b.'s should also be subject to the laws (A·14).

Now, we have encountered with such a non-integrable constraint as (6·13)'. As the P.b.'s are expressed in terms of derivatives, we can also define P.b.'s with respect to χ_σ as follows :

$$[p_\mu^\beta, \chi_\sigma] = g_{\mu\nu}(\mathbf{x}_i, \alpha) A_\sigma^\nu(\mathbf{x}_i), \quad (\text{A} \cdot 15)$$

but it destroys Jacobi's identity. In fact (6·13)' means that any variation of χ_σ should vanish. It is imperative, therefore, that the generalized P.b. to be defined shortly would satisfy both (A·14) and (6·13). Such a generalized P.b. (G.P.B) is defined by the formula

$$[\xi, \eta]^* \equiv [\xi, \eta] + [\xi, \phi_\sigma] \Gamma^{\sigma\rho} [\chi_\rho, \eta] - [\xi, \chi_\rho] \Gamma^{\rho\sigma} [\phi_\sigma, \eta] \quad (\text{A} \cdot 16)$$

equivalent to Dirac's one, and (A·16) gives

$$\begin{aligned} [\xi, \chi_\rho]^* &\equiv [\xi, \chi_\rho] - [\xi, \chi_\tau] \Gamma^{\tau\sigma} [\phi_\sigma, \chi_\rho] \\ &= [\xi, \chi_\rho] - [\xi, \chi_\sigma] \delta_\rho^\sigma = 0, \end{aligned} \quad (\text{A} \cdot 17)$$

because of $[\phi_\sigma, \chi_\rho] = I'_{\sigma\rho}$. There is no ambiguity in this extension of P.b. to include general affine co-ordinates χ_σ because no terms of the second derivatives of χ_σ appear in its definition. The laws (A·14) are easily verified.

We have then Poisson's equations of motion in the modified type

$$d\xi/dt = [H, \xi]^*. \quad (\text{A} \cdot 18)$$

They are equivalent to those which conventionally constructed by taking into account the constraint as Dirac already proved.

In the quantum theory we may interpret the newly defined G.P.B. as the modified commutation relation of operators. This alters the differentiation operation into

$$(\partial/\partial x_i^\mu)^* \equiv \partial/\partial x_i^\mu - g_{\mu\nu}(\mathbf{x}_i; \alpha) A_\sigma^\nu(\mathbf{x}_i) \Gamma^{\sigma\rho} \lambda_\rho; \quad (\text{A} \cdot 19)$$

for instance, in the case of c.m.,

$$(\partial/\partial x_i)^* \equiv \partial/\partial x_i - (1/N) \sum_j \partial/\partial x_j.$$

It is just the outcome of universal substitution of derivations in § 2.

References

- 1) F. Bloch, *Helv. Phys. Acta* **7** (1934), 375.
- 2) M. R. Schafroth, *Nuov. Cim.* **11** (1954), 53.
- 3) T. Kinoshita and Y. Nambu, *Phys. Rev.* **94** (1954), 598.
- 4) H. A. Tolhoek, *Physica* **21** (1954), 1.
- 5) T. Nishiyama, *Prog. Theor. Phys.* **12** (1953), 245; **14** (1955), 37.
- 6) D. R. Inglis, *Phys. Rev.* **96** (1954), 1059; **97** (1955), 701.
- 7) F. Coester, *Phys. Rev.* **97** (1955), 170.
- 8) D. Hill and J. A. Wheeler, *Phys. Rev.* **89** (1953), 1102.
- 9) H. J. Lipkin, A. De Shalit and I. Talmi, *Nuov. Cim.* **12** (1955), 773.
- 10) T. Marumori, J. Yukawa and R. Tanaka, *Prog. Theor. Phys.* **13** (1955), 442.
- 11) T. Miyazima, *Prog. Theor. Phys.* **15** (1956), 255.
- 12) D. Bohm and D. Pines, *Phys. Rev.* **82** (1951), 652; **85** (1952), 338; **92** (1953), 609.
- 13) S. Tomonaga, *Prog. Theor. Phys.* **13** (1955), 467, 482.
- 14) E. Rosenfeld, *Ann. d. Phys.* **5** (1930), 113.
- 15) K. Husimi, *Buturigaku Kôensyû* (in Japanese) **4** (1943), 84.
- 16) K. Husimi and R. Utiyama, *Prog. Theor. Phys.* **5** (1950), 718.
- 17) A. Bohr, "Rotational States of Atomic Nuclei" Dissertation, Kôbenhavn (1954)
- 18) A. Bohr, *Dansk. Med. Fysk.* **26** (1952), No. 14.
- 19) A. Bohr and A. Mottelson, *Dansk. Med. Fysk.* **27** (1954) No. 27.
- 20) For example; E. Cartan, "Sur la structure des groupes de transformation fini et continus" Thèse (1984);—See also Appendix I.
- 21) R. Utiyama, *Phys. Rev.* **101** (1956), 1597.
- 22) C. N. Yang and R. L. Mills, *Phys. Rev.* **96** (1954), 191.
- 23) P.A.M. Dirac, *Canad. Jour. Math.* **2** (1951), 129

Nuclear Magnetic Relaxation in Antiferromagnetics

Tôru MORIYA*

Department of Physics, Osaka University, Osaka

(Received March 16, 1956)

The mechanism of the thermal relaxation of nuclear spins related with the nuclear magnetic resonance in antiferromagnetics is studied theoretically. The predominant mechanism is provided by the fluctuating local magnetic field at the nucleus coming from the electron spins which flip from time to time due to exchange interactions among them. At low temperatures this mechanism can be described in terms of the inelastic scatterings of spin waves by nuclear magnetic moments, and the thermal transition probability can be calculated from this point of view. At high temperatures the relaxation time is obtained from the local field spectra which are calculated with the use of the high temperature approximation assuming a Gaussian distribution.

The calculated relaxation time of protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ agrees well with the experiment both in order of magnitude and in the qualitative nature of its temperature dependence. However, for MnF_2 , whose nuclear resonance has not yet been observed in the range $300\text{--}1.5^\circ\text{K}$, the calculated relaxation time is always longer than 10^{-5} sec and it becomes rapidly longer as temperature is lowered. This contradicts with the interpretation of Bloembergen and Poulis who accounted for the absence of the resonance by assuming that T_1 is shorter than 10^{-6} sec. As a possible origin of the absence of the resonance in the frequency range of their observation, effect of the hyperfine interaction is suggested.

§ 1. Introduction

Nuclear magnetic resonance in antiferromagnetics was first observed by Poulis and Hardeman¹⁾ for protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$. From the resonance frequencies observed at various temperatures under varieties of external magnetic fields, they obtained interesting informations on the antiferromagnetic properties of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ ²⁾, i.e., the spin superstructure, the positions of the protons, the temperature dependence of the sublattice magnetization, etc. Prior to this, Bloembergen and Poulis³⁾ attempted an experiment on fluorine nuclear resonance in MnF_2 but they could not detect it at temperatures between 300°K and 1.5°K . They attributed the absence of this nuclear resonance to a very short relaxation time ($T_1 < 10^{-6}$ sec). Recently, Hardeman⁴⁾ made a measurement of the relaxation time of protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$. According to his result the relaxation time depends strongly on temperature, namely it increases rapidly as temperature is lowered.

In this paper the writer will develop a theory of the nuclear relaxation in antiferromagnetics and apply it to the case of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ and MnF_2 , and compare it with the corresponding experiments.

*) Now at Tokyo Metropolitan University, Tokyo.

In an antiferromagnetic crystal, nuclear spins are subjected to a strong magnetic field produced by the antiferromagnetically arranged electron spins. The average internal field together with the external field determine the Zeeman splitting of the nuclear spin states, while the fluctuation with time of the internal field induces transitions between nuclear Zeeman levels, which provides the mechanism of the thermal relaxation of the nuclear spins. The thermal fluctuation of the electron spins is produced by exchange interactions among electron spins as well as by lattice vibrations. The exchange interactions make spins flip from time to time and this gives rise to large fluctuations of the local magnetic field. The lattice vibrations, however, can give only small fluctuations through the change of the coordinates of the magnetic ions relative to the nucleus. At least for substances whose Curie temperature is far lower than their Debye temperature, such as $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, the thermal relaxation will mainly be provided by the exchange interactions. In fact, the specific heat data of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ ⁵ show that the degrees of freedom of the lattice vibrations are almost dead below its Curie temperature. We shall therefore confine ourselves mainly to the effect of the exchange interactions.

At low temperatures an exchange-coupled electron spin system can well be described by the spin wave approximation. A nuclear magnetic moment scatters spin waves by its dipolar field. There are inelastic scatterings which accompany simultaneous flippings of the nuclear spins, thus providing the nuclear relaxation mechanism. A calculation based on this picture is given in § 2. There is another standpoint which consists in looking for such Fourier components of the fluctuating local magnetic field that correspond to the nuclear Larmor frequency. Such a type of calculation is given in § 3; with the spin wave approximation it naturally leads to the same conclusion as that of § 2, while with the moment expansion method we can calculate the relaxation time above the Curie point. In § 4 we at first apply the theory to the protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ and compare it with the experiment. Then, another application of the theory is made to the fluorine nuclei in MnF_2 , in which case we also estimate a minor contribution from the lattice vibrations, and discussions on the absence of the nuclear resonance in MnF_2 are given referring to a possible hyperfine interaction.

§ 2. Thermal relaxation by inelastic scatterings of spin waves

We shall adopt the two sublattice model for the antiferromagnet, taking account of only nearest neighbour exchange interactions. The magnetic anisotropy is assumed to be uniaxial. A nucleus whose relaxation is now being considered is assumed not to belong to the magnetic ion which itself is regarded as a point dipole. The starting Hamiltonian of the electron spin system and that of the nuclear spin are expressed as follows:

$$\mathfrak{H}_e = J \sum_{\langle jk \rangle} \mathbf{S}_j \cdot \mathbf{S}_k - K \left(\sum_j S_{jx}^2 + \sum_k S_{kx}^2 \right) - \hbar \gamma_e \mathbf{H}_0 \cdot \left(\sum_j \mathbf{S}_j + \sum_k \mathbf{S}_k \right), \quad (2.1)$$

$$\begin{aligned} \mathfrak{H}_N &= -\hbar \gamma_N \mathbf{I} \cdot \left[\mathbf{H}_0 - \hbar \gamma_e \sum_l \mathbf{r}_l^{-3} \{ \langle \mathbf{S}_l \rangle - 3 \mathbf{r}_l (\mathbf{r}_l \cdot \langle \mathbf{S}_l \rangle) / r_l^2 \} \right] \\ &= -\hbar \gamma_N \mathbf{I} \cdot \mathbf{H}, \end{aligned} \quad (2.2)$$

where J is the exchange integral times -2 , K the anisotropy constant, \mathbf{H}_0 the external magnetic field, \mathbf{S}_j and \mathbf{S}_k the spin operators on the plus and the minus sublattices respectively, $\langle \mathbf{S}_l \rangle$ the thermal average of the l -th electron spin, \mathbf{r}_l the position vector of the l -th electron spin referring to the origin at the nucleus, r_l its length, γ_e and γ_N the gyro-magnetic ratios of the electronic and nuclear spins, \mathbf{H} the average local field at the nucleus, and $\sum_{\langle jk \rangle}$ means the summation over the nearest neighboring pairs.

The dipole-dipole interaction between the nuclear spin and the fluctuations of the electron spins is taken as a perturbation :

$$\mathfrak{H}' = \gamma_e \gamma_N \hbar^2 \sum_l r_l^{-3} (\mathbf{I} \cdot \delta \mathbf{S}_l - 3 \mathbf{I} \cdot \frac{\mathbf{r}_l}{r_l^3} \cdot \delta \mathbf{S}_l), \quad (2.3)$$

where

$$\delta \mathbf{S}_l = \mathbf{S}_l - \langle \mathbf{S}_l \rangle. \quad (2.4)$$

The spin wave treatment of our model of the antiferromagnet has been given by several authors.^{6,7)} The result is that when there is no external field one has the following spin wave Hamiltonian :

$$\mathfrak{H}_0 = \sum_{\sigma} \hbar \omega_{\sigma} (n_{1\sigma} + n_{2\sigma}), \quad (2.5)$$

with

$$\hbar \omega_{\sigma} = z J S \{ (1 + \alpha)^2 - \gamma_{\sigma}^2 \}^{1/2}, \quad (2.6)$$

where $n_{1\sigma}$ and $n_{2\sigma}$ denote the number operators of the two degenerate spin waves with equal wave number vector σ , z the number of nearest neighbours, and α and γ_{σ} are defined by

$$\alpha = 2K/zJ, \quad (2.7)$$

$$\gamma_{\sigma} = 1/z \cdot \sum_{\rho} \exp(i\rho \cdot \sigma), \quad (2.8)$$

ρ denoting the relative position vector of any one nearest neighbour and the summation extending over all the nearest neighbours. When there is an external field, it lifts the two-fold degeneracy of the σ -waves but the relaxation rate is not much affected by it, as we shall see below, so that we omit henceforth the external field in treating the spin waves.

The perturbing Hamiltonian \mathfrak{H}' can be expressed in terms of annihilation and creation operators of spin waves. Linear terms with respect to these operators give rise to nuclear spin transitions accompanied by an absorption or emission of one spin wave. This will be called direct processes; only those spin waves whose energy corresponds exactly to the nuclear Larmor frequency can take part in these processes. However, as is easily seen from (2.6), such a spin wave does not usually exist because of the existence of the anisotropy field. Even if there is an external field, which makes some of the spin wave frequencies shift to the lower side, there is no participant in the direct processes unless the external field come very close to the critical field. The dominant effect comes from Raman processes, which are related to bilinear terms with respect to the creation and annihilation

operators. In a Raman process a nuclear spin flip is accompanied by simultaneous emission of one spin wave and absorption of another, while the total energy is conserved. Almost all modes of spin waves can take part in these processes so that a large transition probability is expected from them. We shall treat only these processes.

The perturbing Hamiltonian (2.3) is expressible in terms of the spin deviation operators, a_j^* , a_j , b_k^* , b_k , defined by

$$\begin{aligned} S_{jx} + iS_{jy} &= (2S)^{1/2} [1 - (a_j^* a_j / 2S)]^{1/2} a_j, \\ S_{jx} - iS_{jy} &= (2S)^{1/2} a_j^* [1 - (a_j^* a_j / 2S)]^{1/2}, \\ S_{jz} &= S - a_j^* a_j; \\ S_{kx} + iS_{ky} &= (2S)^{1/2} b_k^* [1 - (b_k^* b_k / 2S)]^{1/2}, \\ S_{kx} - iS_{ky} &= (2S)^{1/2} [1 - (b_k^* b_k / 2S)]^{1/2} b_k, \\ S_{kz} &= -S + b_k^* b_k. \end{aligned}$$

Taking only the second order terms with respect to these operators, we have

$$\mathfrak{H}' = -\gamma_c \gamma_N \hbar^2 \sum_l f_l n_l, \quad (2.9)$$

with

$$n_j = a_j^* a_j, \quad n_k = b_k^* b_k, \quad (2.10)$$

$$f_l = r_l^{-3} [I_z - 3(\mathbf{I} \cdot \mathbf{r}_l) z_l / r_l^2], \quad l = j \text{ or } k. \quad (2.11)$$

The spin deviation operators are expressed in terms of the spin wave operators, $a_{1\sigma}^*$, $a_{1\sigma}$, $a_{2\sigma}^*$, $a_{2\sigma}$, as follows:

$$\begin{aligned} a_j &= (2/N)^{1/2} \sum_{\sigma>0} 1 / \sqrt{2A_\sigma B_\sigma} \cdot [(A_\sigma + B_\sigma) \{a_{1\sigma} \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_j) + a_{1-\sigma} \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_j)\} \\ &\quad - (A_\sigma - B_\sigma) \{a_{2\sigma}^* \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_j) + a_{2-\sigma}^* \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_j)\}], \\ a_j^* &= (2/N)^{1/2} \sum_{\sigma>0} 1 / \sqrt{2A_\sigma B_\sigma} \cdot [(A_\sigma + B_\sigma) \{a_{1\sigma}^* \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_j) + a_{1-\sigma}^* \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_j)\} \\ &\quad - (A_\sigma - B_\sigma) \{a_{2\sigma} \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_j) + a_{2-\sigma} \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_j)\}], \\ b_k &= (2/N)^{1/2} \sum_{\sigma>0} 1 / \sqrt{2A_\sigma B_\sigma} \cdot [-(A_\sigma - B_\sigma) \{a_{1\sigma}^* \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_k) + a_{1-\sigma}^* \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_k)\} \\ &\quad + (A_\sigma + B_\sigma) \{a_{2\sigma} \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_k) + a_{2-\sigma} \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_k)\}], \\ b_k^* &= (2/N)^{1/2} \sum_{\sigma>0} 1 / \sqrt{2A_\sigma B_\sigma} \cdot [-(A_\sigma - B_\sigma) \{a_{1\sigma} \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_k) + a_{1-\sigma} \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_k)\} \\ &\quad + (A_\sigma + B_\sigma) \{a_{2\sigma}^* \cos(\boldsymbol{\sigma} \cdot \mathbf{r}_k) + a_{2-\sigma}^* \sin(\boldsymbol{\sigma} \cdot \mathbf{r}_k)\}], \end{aligned} \quad (2.12)$$

with

$$A_\sigma = \sqrt{1 + \alpha + \gamma_\sigma}, \quad B_\sigma = \sqrt{1 + \alpha - \gamma_\sigma},$$

where N is the number of electron spins and the summations are taken over half the σ -

space.

The transition matrix element of a Raman process is

$$(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1; m+1 | \tilde{\zeta}' | n_{\nu\sigma}, n_{\nu'\sigma'}; m) \\ = -\gamma_e \tilde{\gamma}_N \tilde{b}^2 \sum (m+1 | f_l | m) (n_{\nu\sigma}+1, n_{\nu'\sigma'}-1 | n_l | n_{\nu\sigma}, n_{\nu'\sigma'}) \cdot (\nu=1, 2) \quad (2 \cdot 13)$$

The matrix elements of n_l can be calculated by using (2.10) and (2.12). The results are

$$(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1 | n_j | n_{\nu\sigma}, n_{\nu'\sigma'}) \\ = \delta_{\nu\nu'} \left(\frac{2}{N} \right) \frac{(A_\sigma \pm B_\sigma) (A_{\sigma'} \pm B_{\sigma'})}{\sqrt{2A_\sigma B_\sigma} \sqrt{2A_{\sigma'} B_{\sigma'}}} \{ (n_{\nu\sigma}+1) n_{\nu\sigma'} \}^{1/2} \times \\ \times \begin{cases} \cos(\sigma \cdot \mathbf{r}_j) \cos(\sigma' \cdot \mathbf{r}_j) & \text{for } \sigma > 0, \sigma' > 0; \\ \sin(\sigma \cdot \mathbf{r}_j) \sin(\sigma' \cdot \mathbf{r}_j) & \text{for } \sigma < 0, \sigma' < 0; \\ -\cos(\sigma \cdot \mathbf{r}_j) \sin(\sigma' \cdot \mathbf{r}_j) & \text{for } \sigma > 0, \sigma' < 0; \\ -\sin(\sigma \cdot \mathbf{r}_j) \cos(\sigma' \cdot \mathbf{r}_j) & \text{for } \sigma < 0, \sigma' > 0, \end{cases} \quad (2 \cdot 14a)$$

$$(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1 | n_k | n_{\nu\sigma}, n_{\nu'\sigma'}) \\ = -\delta_{\nu\nu'} (2/N) \frac{(A_\sigma \mp B_\sigma) (A_{\sigma'} \mp B_{\sigma'})}{\sqrt{2A_\sigma B_\sigma} \sqrt{2A_{\sigma'} B_{\sigma'}}} \{ (n_{\nu\sigma}+1) n_{\nu\sigma'} \}^{1/2} \times \\ \times \begin{cases} \cos(\sigma \cdot \mathbf{r}_k) \cos(\sigma' \cdot \mathbf{r}_k) & \text{for } \sigma > 0, \sigma' > 0; \\ \sin(\sigma \cdot \mathbf{r}_k) \sin(\sigma' \cdot \mathbf{r}_k) & \text{for } \sigma < 0, \sigma' < 0; \\ -\cos(\sigma \cdot \mathbf{r}_k) \sin(\sigma' \cdot \mathbf{r}_k) & \text{for } \sigma > 0, \sigma' < 0; \\ -\sin(\sigma \cdot \mathbf{r}_k) \cos(\sigma' \cdot \mathbf{r}_k) & \text{for } \sigma < 0, \sigma' > 0, \end{cases} \quad (2 \cdot 14b)$$

where ν and ν' take 1 or 2, and the upper one of the double sign is to be taken for $\nu=\nu'=1$ and the lower one for $\nu=\nu'=2$. The matrix elements of f_l depend on the direction of the local magnetic field at the nucleus. We take this direction as ζ -axis and define $\hat{\xi} \eta \zeta$ -coordinate system. The direction cosines of $\hat{\xi}$, η and ζ axes are denoted by $(\alpha_1, \beta_1, \gamma_1)$, $(\alpha_2, \beta_2, \gamma_2)$ and (α, β, γ) ; similarly those of the position vector \mathbf{r}_l by $(\alpha_l, \beta_l, \gamma_l)$. Then f_l can be expressed as

$$f_l = r_l^{-3} [(1-3\gamma_l^2) (\gamma_l I_\xi + \gamma_2 I_\eta + \gamma_l I_\zeta) - 3\gamma_l \alpha_l (\alpha_l I_\xi + \alpha_2 I_\eta + \alpha_l I_\zeta) \\ - 3\gamma_l \beta_l (\beta_l I_\xi + \beta_2 I_\eta + \beta_l I_\zeta)]. \quad (2 \cdot 15)$$

As the nuclear spin is quantized along the ζ -axis, the transition matrix element of f_l becomes

$$(m+1 | f_l | m) = (1/2) \gamma_l^{-3} [(1-3\gamma_l^2) (\gamma_1 - i\gamma_2) - 3\gamma_l \alpha_l (\alpha_1 - i\alpha_2) - 3\gamma_l \beta_l (\beta_1 - i\beta_2)] \\ \times \sqrt{(I-m)(I+m+1)}.$$

This gives

$$|(m+1|f_l|m)|^2 = (1/4) r_l^{-6} F_l(\alpha\beta\gamma) (I-m) (I+m+1), \quad (2.16)$$

with

$$F_l(\alpha\beta\gamma) = (1-6\gamma_l^2) (1-\gamma^2) + 9\gamma_l^2 \{1 - (\alpha_l^2\alpha^2 + \beta_l^2\beta^2 + \gamma_l^2\gamma^2)\} \\ + 6(1-3\gamma_l^2) (\alpha_l\gamma_l\alpha\gamma + \beta_l\gamma_l\beta\gamma) - 18\gamma_l^2 \alpha_l\beta_l\alpha\beta. \quad (2.17)$$

The thermal transition probability of the nuclear spin corresponding to $I_z = m \rightarrow m+1$ is given by

$$W(m \rightarrow m+1) = \frac{2\pi}{\hbar} \sum_{\nu\nu'} \sum_{\sigma\sigma'} \langle |(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1; m+1|\mathfrak{H}'|n_{\nu\sigma}, n_{\nu'\sigma'}; m)|^2 \rangle \\ \times \delta(\hbar\omega_\sigma - \hbar\omega_{\sigma'} - \hbar\gamma_N H), \quad (2.18)$$

where $\langle \rangle$ means the statistical average with respect to the electron spin system; we naturally assume that the electron spin system is in thermal equilibrium. In the above summations, those terms which satisfy both $\nu=\nu'$ and $\sigma=\sigma'$ should be omitted. From (2.13) we have

$$|(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1; m+1|\mathfrak{H}'|n_{\nu\sigma}, n_{\nu'\sigma'}; m)|^2 \\ = \gamma_e^2 \gamma_N^2 \hbar^4 \left[\sum_l |(m+1|f_l|m)|^2 |(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1|n_l|n_{\nu\sigma}, n_{\nu'\sigma'})|^2 \right. \\ \left. + \sum_{l \neq l'} (m|f_l|m+1)(m+1|f_{l'}|m) (n_{\nu\sigma}, n_{\nu'\sigma'}|n_l|n_{\nu\sigma}+1, n_{\nu'\sigma'}-1) \right. \\ \left. \times (n_{\nu\sigma}+1, n_{\nu'\sigma'}-1|n_{l'}|n_{\nu\sigma}, n_{\nu'\sigma'}) \right]. \quad (2.19)$$

In this expression only a few terms make important contributions to the transition probability because of the factor r_l^{-6} . Terms in the second summation represent the effect of interference of two components of the fluctuating local field coming from different two electron spins, and since their values take positive as well as negative signs we can expect that they contribute negligibly to the transition probability. So we omit them.

Inserting (2.19) and (2.16) into (2.18) and replacing the summation over σ and σ' by an integral, we get

$$W(m \rightarrow m+1) = \frac{\pi}{2\hbar} \gamma_e^2 \gamma_N^2 \hbar^4 \sum_l r_l^{-6} F_l(\alpha\beta\gamma) (I-m) (I+m+1) \\ \times \sum_{\nu} \frac{V^2}{(2\pi)^6} \iint d\sigma d\sigma' \langle |(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1|n_l|n_{\nu\sigma}, n_{\nu'\sigma'})|^2 \rangle \delta(\hbar\omega_\sigma - \hbar\omega_{\sigma'}), \quad (2.20)$$

where V is the volume of the specimen. The nuclear Larmor frequency is here neglected in the argument of the δ -function since it is small compared to the spin wave frequencies. From (2.14) we can calculate the integrand of (2.20); taking the average over the four types of Raman processes appearing in (2.14a) and (2.14b), we get

$$\sum_{\nu} (1/4) \langle |(n_{\nu\sigma}+1, n_{\nu'\sigma'}-1|n_l|n_{\nu\sigma}, n_{\nu'\sigma'})|^2 + |(n_{\nu-\sigma}+1, n_{\nu\sigma'}-1|n_l|n_{\nu-\sigma}, n_{\nu\sigma'})|^2 \\ + |(n_{\nu\sigma}+1, n_{\nu-\sigma'}-1|n_l|n_{\nu\sigma}, n_{\nu-\sigma'})|^2 + |(n_{\nu-\sigma}+1, n_{\nu\sigma'}-1|n_l|n_{\nu-\sigma}, n_{\nu\sigma'})|^2 \rangle$$

$$= \left(\frac{2}{N}\right)^2 \frac{1}{2} \langle n_\sigma + 1 \rangle \langle n_{\sigma'} \rangle \left[\frac{(1+\alpha)^2}{\sqrt{(1+\alpha)^2 - \gamma_\sigma^2} \sqrt{(1+\alpha)^2 - \gamma_{\sigma'}^2}} + 1 \right], \quad (2 \cdot 21)$$

where we used the relation $\langle n_{1\sigma} \rangle = \langle n_{1-\sigma} \rangle = \langle n_{2\sigma} \rangle = \langle n_{2-\sigma} \rangle$ and denoted these quantities as $\langle n_\sigma \rangle$, which is

$$\langle n_\sigma \rangle = (e^{\hbar\omega_\sigma/\kappa T} - 1)^{-1}, \quad (2 \cdot 22)$$

κ being the Boltzmann constant.

Next we define the relative state density by

$$N(\omega) = \frac{2}{N} \frac{V}{(2\pi)^3} \iint \frac{dS}{|\text{grad}_\sigma \omega|} \quad (2 \cdot 23)$$

where grad_σ means the differentiation in the σ -space and the double integral is taken over an energy surface $\omega = \text{const.}$ in the σ -space. $N(\omega)$ is normalized to 1:

$$\int_0^{\omega_m} N(\omega) d\omega = 1, \quad (2 \cdot 24)$$

ω_m being the maximum frequency of the spin waves. By using (2.21), (2.22) and (2.23), transition probability (2.20) can be written as

$$\begin{aligned} W(m \rightarrow m+1) &= \frac{1}{4} \pi \gamma_e^2 \gamma_N^2 \hbar^2 (I-m)(I+m+1) \sum_l r_l^{-6} F_l(\alpha\beta\gamma) \\ &\times \int_0^{\omega_m} \frac{e^{\hbar\omega/\kappa T}}{(e^{\hbar\omega/\kappa T} - 1)^2} \left\{ \left(\frac{(1+\alpha)zJS}{\hbar\omega} \right)^2 + 1 \right\} \{N(\omega)\}^2 d\omega. \end{aligned} \quad (2 \cdot 25)$$

Defining characteristic temperature θ and reduced temperature T^* by

$$\theta = \hbar\omega_m/\kappa, \quad T^* = T/\theta, \quad (2 \cdot 26)$$

we get the final expression for the thermal transition probability:

$$\begin{aligned} W(m \rightarrow m+1) &= \frac{1}{4} \pi \gamma_e^2 \gamma_N^2 \hbar^2 (I-m)(I+m+1) \sum_l r_l^{-6} F_l(\alpha\beta\gamma) \\ &\times \omega_m T^* \int_0^{1/T^*} \frac{e^x}{(e^x - 1)^2} \left\{ \left(\frac{(1+\alpha)zJS}{\kappa\theta} \right)^2 \frac{1}{T^{*2}} \frac{1}{x^2} + 1 \right\} \{N(\omega_m T^* x)\}^2 dx. \end{aligned} \quad (2 \cdot 27)$$

We have now only to calculate $N(\omega)$ from (2.23) and (2.6). As it is troublesome to carry out the exact calculation of $N(\omega)$, we shall content ourselves here with taking the long wave approximation:

$$\hbar\omega_\sigma = zJS \sqrt{(\gamma n^{-1/3} \sigma)^2 + 2\alpha}, \quad (2 \cdot 28)$$

where n is the number density of spins and η a geometrical factor which depends on the spin super-structure: $\eta = 3^{-1/2}$ for NaCl type and $\eta = 2^{-2/3}$ for CsCl type. Since the expression (2.28) depends only on the magnitude of σ , we can restrict the wave number space to the inside of a sphere which includes just $N/2$ modes. Then the maximum frequency is determined as

$$\hbar\omega_m = zJS \sqrt{\{\gamma(3\pi^2)^{1/3}\}^2 + 2\alpha} \simeq (3\pi^2)^{1/3} \gamma zJS = \kappa\theta. \quad (2.29)$$

Further we shall define a temperature T_{AE} and a frequency ω_{AE} which measure the effect of the interplay between the anisotropy and exchange energies.⁸⁾:

$$\kappa T_{AE} = \hbar\omega_{AE} = \sqrt{2\alpha} zJS = 2S \sqrt{zJK}. \quad (2.30)$$

Then the state density becomes

$$\begin{aligned} N(\omega) &= (3/\omega_m^3) \omega \sqrt{\omega^2 - \omega_{AE}^2} \quad \text{for } \omega_m > \omega > \omega_{AE}, \\ &= 0 \quad \text{for } \omega < \omega_{AE}. \end{aligned} \quad (2.31)$$

Inserting (2.31) into (2.27) we get the following expression for the transition probability:

$$\begin{aligned} W(m \rightarrow m+1) &= \frac{9}{4} \frac{\pi \gamma_e^2 \gamma_N^2 \hbar^2}{\kappa \theta} (I-m)(I+m+1) \sum_l r_l^{-6} F_l(\alpha \beta \gamma) \\ &\quad \times [D_1(T^*) + D_2(T^*) + D_3(T^*)], \end{aligned} \quad (2.32)$$

where

$$\begin{aligned} D_1(T^*) &= [\{(1+\alpha)/(3\pi^2)^{1/3}\gamma\}^2 - (T_{AE}/\theta)^2] T^{*3} \{\mathfrak{F}_2(1/T^*) - \mathfrak{F}_2(T_{AE}/T)\}, \\ D_2(T^*) &= T^{*5} \{\mathfrak{F}_4(1/T^*) - \mathfrak{F}_4(T_{AE}/T)\}, \end{aligned} \quad (2.33)$$

$$D_3(T^*) = \{(1+\alpha)/(3\pi^2)^{1/3}\gamma\}^2 (T_{AE}/\theta)^2 T^* \{(e^{1/T^*} - 1)^{-1} - (e^{T_{AE}/T} - 1)^{-1}\},$$

$$\mathfrak{F}_n(x) = \int_0^x \frac{z^n dz}{(e^z - 1)(1 - e^{-z})}. \quad (2.34)$$

Numerical values of $\mathfrak{F}_2(x)$ and $\mathfrak{F}_4(x)$ are listed in Table I. The thermal relaxation time T_1 is calculated from the transition probability with the use of the relation⁹⁾

$$T_1 = (I-m)(I+m+1)/2W(m \rightarrow m+1). \quad (2.35)$$

Table I.

| x | $\mathfrak{F}_2(1/x)$ | $\mathfrak{F}_4(1/x)$ | x | $\mathfrak{F}_2(1/x)$ | $\mathfrak{F}_4(1/x)$ |
|------|-----------------------|-----------------------|-----|-----------------------|-----------------------|
| 0 | 3.29 | 25.98 | 0.6 | 1.55 | 1.35 |
| 0.05 | 3.29 | 25.98 | 0.8 | 1.20 | 0.603 |
| 0.1 | 3.28 | 25.27 | 1.0 | 0.973 | 0.317 |
| 0.2 | 3.04 | 15.36 | 2.0 | 0.497 | 0.0412 |
| 0.3 | 2.56 | 7.50 | 3.0 | 0.332 | 0.0123 |
| 0.4 | 2.14 | 3.88 | 4.0 | 0.250 | 0.00519 |

§ 3. Fluctuations of the local magnetic field— T_1 above T_c

In this section we shall calculate the thermal relaxation rate by studying the Fourier spectra of the fluctuating local magnetic field. For this purpose it is sufficient to calculate

the correlation function of the local field, which is the Fourier transform of the local field spectra. We suppose that the fluctuation arises due mainly to the exchange interactions and omit the effect of the lattice vibrations. For the quantum mechanical treatment of the correlation function, we can make use of Kubo-Tomita's theory.¹⁰⁾

Using the notation of the previous section, T_1 is given by the following formula (see Appendix):

$$\frac{1}{T_1} = \frac{1}{2} \gamma_e^2 \gamma_N^2 \hbar^2 \sum_i r_i^{-6} [F_i(\alpha\beta\gamma) \int_{-\infty}^{\infty} dt \cos\omega_0 t \langle \delta S_{iz}(t) \delta S_{iz} \rangle + F_i'(\alpha\beta\gamma) \int_{-\infty}^{\infty} dt \cos\omega_0 t \frac{1}{4} (\langle \delta S_i^+(t) \delta S_i^- \rangle + \langle \delta S_i^-(t) \delta S_i^+ \rangle)], \quad (3.1)$$

with

$$\begin{aligned} \delta S_i^{\pm} &= \delta S_{iz} \pm i \delta S_{iy}, \\ \delta S_i(t) &= e^{it\hat{S}_i e / \hbar} \delta S_i e^{-it\hat{S}_i e / \hbar} \end{aligned} \quad (3.2)$$

where ω_0 is the nuclear Larmor frequency and $\langle \rangle$ means the statistical average:

$$\langle Q \rangle = \text{Tr}[e^{-\beta_0 \hat{S}_i e} Q] / \text{Tr}[e^{-\beta_0 \hat{S}_i e}], \quad \beta_0 = 1/\kappa T. \quad (3.3)$$

$F_i(\alpha\beta\gamma)$ has already been defined by (2.17) and $F_i'(\alpha\beta\gamma)$ is

$$\begin{aligned} F_i'(\alpha\beta\gamma) &= 1 + \gamma^2 + 9(1 - \gamma_i^2) [1 - (\alpha_i \alpha + \beta_i \beta + \gamma_i \gamma)^2] \\ &\quad - 6[1 - \gamma_i^2 - (\alpha_i \alpha + \beta_i \beta)^2 - \alpha_i \gamma_i \alpha \gamma - \beta_i \gamma_i \beta \gamma]. \end{aligned} \quad (3.4)$$

In deriving formula (3.1), we omitted, as in the previous section, the correlations between different two electron spins.

The exact calculation of the correlation function is very difficult. At present we know no other ways than to make the spin wave approximation to calculate the whole functional form of the correlation function. However, we can develop it in a power series in t , where the coefficient of $t^n/n!$ is the n -th moment of the local field spectra, and even when we cannot actually carry out the calculation of the whole series we can sometimes assume a reasonable functional form of the spectra by a physical insight and determine the parameters contained in it by calculating a few lower order terms of the series. Thus, when a Gaussian distribution is assumed, it is sufficient to calculate only the first and second moments. In our case, the Gaussian assumption seems to be reasonable at temperatures high enough above the Curie point. Therefore, we shall make use of the high temperature approximation, or the expansion in powers of $\beta_0 J$, to calculate the first and second moments. Other statistical approximation methods¹¹⁾ can also be utilized to calculate the moments, though it is difficult to assume a reasonable functional form of the local field spectra or the correlation function in the neighbourhood of or below the Curie temperature. Actually the spin wave calculation shows that the local field spectra has no simple form; there is a high frequency part of the spectra which does not vanish even at 0°K . We calculate below the correlation function by the spin wave approxima-

tion and by the high temperature approximation.

A) Spin wave approximation

As seen in the previous section, ∂S is expressed as follows:

$$\begin{aligned}\partial S_{jz} &= -a_j^* a_j + \langle a_j^* a_j \rangle, \\ \partial S_j^+ &= (2S)^{1/2} a_j, \\ \partial S_j^- &= (2S)^{1/2} a_j^*,\end{aligned}\tag{3.5}$$

neglecting terms higher than those quadratic in the spin deviation operators. We have similar expressions for suffix k with b_k and b_k^* . a_j , a_j^* and b_k , b_k^* are expressed by (2.12) in terms of the spin wave operators which satisfy

$$\begin{aligned}e^{it\tilde{\Phi}_e/\hbar} a_{\nu\sigma} e^{-it\tilde{\Phi}_e/\hbar} &= e^{-i\omega_\sigma t} a_{\nu\sigma}, \\ e^{it\tilde{\Phi}_e/\hbar} a_{\nu\sigma}^* e^{-it\tilde{\Phi}_e/\hbar} &= e^{i\omega_\sigma t} a_{\nu\sigma}^*.\end{aligned}\tag{3.6}$$

From (3.2), (3.3), (3.5), (3.6) and (2.12) we can directly calculate the correlation functions. We get

$$\begin{aligned}\langle \partial S_{jz}(t) \partial S_{jz} \rangle &= \left(\frac{2}{N}\right)^2 \sum_{\sigma > 0} \sum_{\sigma' > 0} \left[\left(1 + \frac{1+\alpha}{\sqrt{(1+\alpha)^2 - \gamma_\sigma^2}}\right) \left(1 + \frac{1+\alpha}{\sqrt{(1+\alpha)^2 - \gamma_{\sigma'}^2}}\right) \right. \\ &\quad \times \langle n_{1\sigma} \rangle \langle n_{1\sigma'} + 1 \rangle e^{i(\omega_\sigma - \omega_{\sigma'})t} \\ &\quad \left. + \left(1 - \frac{1+\alpha}{\sqrt{(1+\alpha)^2 - \gamma_\sigma^2}}\right) \left(1 - \frac{1+\alpha}{\sqrt{(1+\alpha)^2 - \gamma_{\sigma'}^2}}\right) \langle n_{2\sigma} + 1 \rangle \langle n_{2\sigma'} \rangle e^{-i(\omega_\sigma - \omega_{\sigma'})t} \right] \\ &+ (\text{high frequency terms}),\end{aligned}\tag{3.7}$$

where the high frequency terms arise from the quadratic terms with respect to the spin wave annihilation operators only or those with respect to the spin wave creation operators only which appear when (2.12) is substituted in $a_j^* a_j$ and $b_k^* b$ and include no adequate frequency components to the nuclear relaxation. Other correlation functions than (3.7), i.e., $\langle \partial S_j^+(t) \partial S_j^- \rangle$ and $\langle \partial S_j^-(t) \partial S_j^+ \rangle$, have just the same frequency components as the spin wave spectra and accordingly have no adequate frequency components to the nuclear relaxation. Thus we get the following expression for T_1 :

$$\begin{aligned}\frac{1}{T_1} &= 2\pi \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l r_l^{-6} F_l(\alpha \beta \gamma) (2/N)^2 \\ &\quad \times \sum_{\substack{\sigma > 0, \sigma' > 0 \\ \sigma \neq \sigma'}} \left\{ 1 + \frac{(1+\alpha)^2}{\sqrt{(1+\alpha)^2 - \gamma_\sigma^2} \sqrt{(1+\alpha)^2 - \gamma_{\sigma'}^2}} \right\} \frac{e^{\hbar \omega_\sigma / KT}}{(e^{\hbar \omega_\sigma / KT} - 1)^2} \delta(\omega_\sigma - \omega_{\sigma'}).\end{aligned}$$

Using (2.6) and (2.23) we get

$$\frac{1}{T_1} = \frac{1}{2} \pi \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l r_l^{-6} F_l(\alpha \beta \gamma) \int \left\{ 1 + ((1+\alpha) z J S / \hbar \omega)^2 \right\} \frac{e^{\hbar \omega / \kappa T}}{(e^{\hbar \omega / \kappa T} - 1)^2} \{N(\omega)\}^2 d\omega.\tag{3.8}$$

This is exactly the same result as that we derived in the previous section.

B) High temperature approximation

Now we wish to calculate the correlation function as a power series in t , i.e.,

$$\begin{aligned}\langle S_i(t) S_i(0) \rangle &= \langle e^{it\tilde{\mathfrak{H}}_e/\hbar} S_i e^{-it\tilde{\mathfrak{H}}_e/\hbar} S_i \rangle \\ &= \langle S_i S_i \rangle + (it/\hbar) \langle [\tilde{\mathfrak{H}}_e, S_i] S_i \rangle + 1/2 (it/\hbar)^2 \langle [\tilde{\mathfrak{H}}_e, S_i] [\tilde{\mathfrak{H}}_e, S_i] \rangle + \dots\end{aligned}\quad (3.9)$$

Each coefficient in this series expansion can be calculated as a power series of β_0 by the well-known procedure. After straightforward but lengthy calculations we get the following result up to the second power of β_0 :

$$\begin{aligned}\langle S_{iz}(t) S_{iz} \rangle &= \frac{1}{2} \langle S_i^+(t) S_i^- \rangle = \frac{1}{2} \langle S_i^-(t) S_i^+ \rangle \\ &= \frac{1}{3} S(S+1) [1 - it\beta_0\omega_e\eta_1(1 - \beta_0 J\eta_2) \\ &\quad - \frac{1}{2} t^2 \omega_e^2 \{1 - \beta_0 J\eta_2 + \frac{1}{2} (\beta_0 J)^2 \eta_3 - \dots\} + \dots],\end{aligned}\quad (3.10)$$

where

$$\omega_e^2 = \frac{2}{3} \frac{J^2}{\hbar^2} z S(S+1), \quad (3.11)$$

$$\begin{aligned}\eta_1 &= \sqrt{(z/6) S(S+1)}, \quad \eta_2 = \frac{1}{3} z' S(S+1) - \frac{1}{4}, \\ \eta_3 &= \frac{1}{5} [S(S+1) + \frac{1}{2}]^2 + \frac{1}{6} S(S+1) (z - 2z' - 1) \\ &\quad + (4z''/9 - 1/3) [S(S+1)]^2,\end{aligned}\quad (3.12)$$

In these expressions we have neglected the small effect of the external magnetic field compared with that of the exchange; z means the number of nearest neighbours of each spin, z' the number of nearest neighbours common to two neighbouring spins and z'' is a certain more complicated geometrical factor: their values are

$$\begin{aligned}z=6, \quad z'=0, \quad z''=2 \quad \text{for NaCl type;} \\ z=8, \quad z'=0, \quad z''=6 \quad \text{for CsCl type.}\end{aligned}\quad (3.13)$$

We shall assume a Gaussian distribution for the local field spectra. Then the correlation function can be written

$$\begin{aligned}\langle S_{iz}(t) S_{iz} \rangle &= \frac{1}{3} S(S+1) \exp[-it\beta_0 J\omega_e\eta_1(1 - \beta_0 J\eta_2) \\ &\quad - \frac{1}{2} t^2 \omega_e^2 \{1 - \beta_0 J\eta_2 + \frac{1}{2} (\beta_0 J)^2 (\eta_3 - 2\eta_1^2)\}].\end{aligned}\quad (3.14)$$

Inserting (3.14) into (3.1) the relaxation time is written

$$\frac{1}{T_1} = \frac{1}{T_{1\infty}} [1 + \frac{1}{4} \beta_0 J + \frac{1}{2} (\beta_0 J)^2 (\eta_3 - 2\eta_1^2)]^{-1/2} \exp[-\frac{1}{2} (\beta_0 J)^2 \eta_1^2 (1 + \frac{1}{4} \beta_0 J)], \quad (3.15)$$

with

$$\frac{1}{T_{1\infty}} = \frac{\sqrt{2\pi}}{6} \frac{\gamma_e^2 \gamma_N^2 \hbar^2}{\omega_e} S(S+1) \sum_i \gamma_i^{-6} [F_i(\alpha\beta\gamma) + F_i'(\alpha\beta\gamma)]. \quad (3.16)$$

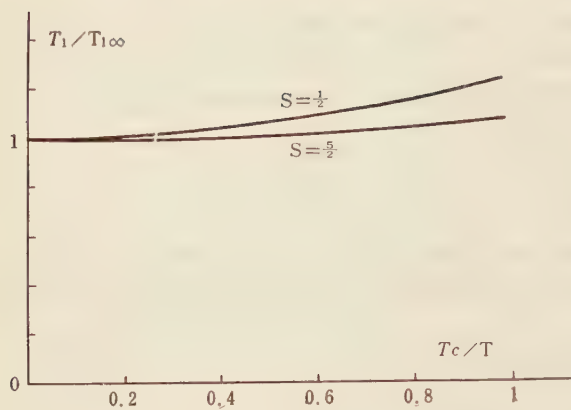


Fig. 1

$T_{1\infty}$ is the limiting value of T_1 for $\beta_0 \rightarrow 0$, or $T \rightarrow \infty$. In this expression we neglected the quantity ω_0/ω_c which is very small compared with $\beta_0 J$ in the temperature region of our interest.

(3.15) shows that T_1 becomes longer as temperature is lowered, which is due to setting in of the short-range order. Fig. 1 shows the temperature dependence of T_1 calculated by using (3.15) for the body centred lattice. The temperature dependence is

not remarkable in this approximation.

§ 4. Comparison with experiment

Hardeman⁴⁾ measured T_1 for the protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ below the Curie temperature, which is the only available data of this kind at the present time. Another existing experiment is that of Bloembergen and Poulis³⁾ for MnF_2 , who, however, failed to detect any nuclear resonance in it. We shall apply our theory to these two substances and discuss the comparison with experiment.

(I) $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$

The crystal structure of this substance has ortho-rhombic symmetry and its spin superstructure is reported to consist of ferromagnetic (001) layers with spins pointing to the $+a$ and $-a$ directions alternately in the c -direction. Though our model of the anti-ferromagnet used in the previous sections is not strictly applicable to this substance, it may still be possible to use it for evaluating the relaxation time if the parameters appearing in the formula could be properly determined from the observed quantities.

The exchange integral is correlated to the perpendicular susceptibility with the use of the molecular field theory by

$$Jz = (N/2) g^2 \mu_B^2 / \chi_{\perp}. \quad (4.1)$$

Taking the average of the experimental values¹²⁾ of χ_{110} , χ_{111} , and χ_{112} , and estimating N from the lattice constants¹³⁾, we get

$$Jz/\kappa = 26.9^\circ\text{K}, \quad (4.2)$$

which gives by (2.29)

$$\theta = 20^\circ\text{K}, \quad (4.3)$$

putting roughly $\eta = 1/2$, since $\eta = 3^{-1/2} = 0.58$ for NaCl structure ($z=6$) and $\eta = 2^{-2/3} = 0.63$ for CsCl structure ($z=8$), while we have here $z=2$, assuming η to change linearly

with z . The anisotropy constants have been estimated from critical field data¹²⁾ or anti-ferromagnetic resonance data¹⁴⁾¹⁵⁾: Taking the average of the two anisotropy constants, one estimates

$$T_{AE}=1.5^{\circ}\text{K}. \quad (4.4)$$

The lattice constants are¹³⁾

$$a=7.38 \text{ \AA}, \quad b=8.04 \text{ \AA}, \quad c=3.72 \text{ \AA}. \quad (4.5)$$

The Cu^{++} ions occupy two positions $(0, 0, 0)$ and $(\frac{1}{2}, \frac{1}{2}, 0)$ in the unit cell. The positions of the protons have been determined by Poulis and Hardeman¹⁾ from their proton resonance data; though these positions are contradictory to the room temperature data given by Itoh and his collaborators¹⁶⁾ and also to the geometry of the free water molecule, we shall here adopt them since small change in the positions does not seriously affect the value of T_1 . Among eight protons in the unit cell we consider one whose coordinates are

$$x=0.98 \text{ \AA}, \quad y=3.05 \text{ \AA}, \quad z=0.92 \text{ \AA}. \quad (4.6)$$

Using these parameters we calculate r_i^{-6} and $F_i(\alpha\beta\gamma)$. Some values of r_i^{-6} are given in Table II; we see that the effect of the neighboring two Cu^{++} ions are dominating. The factor $F_i(\alpha\beta\gamma)$ is of order unity and its directional change is within a factor of about 2. Taking directional average we have

$$\sum_i r_i^{-6} \langle F_i(\alpha\beta\gamma) \rangle = 4 \times 10^{15}. \quad (4.7)$$

With the use of the above numerical values we can calculate T_1 by (2.32) and (2.35), the results of which are shown in Fig. 2 with a solid curve.

Table II.

| position of Cu^{++} | r_i (Å) | r_i^{-6} (cgs) | position of Cu^{++} | r_i (Å) | r_i^{-6} (cgs) |
|------------------------------|-----------|------------------------|------------------------------|-----------|-----------------------|
| $(\frac{1}{2}\frac{1}{2}0)$ | 3.01 | -1.31×10^{15} | (001) | 4.25 | 0.17×10^{15} |
| (000) | 3.33 | 0.73 | $(-\frac{1}{2}\frac{1}{2}0)$ | 4.86 | 0.076 |
| $(\frac{1}{2}\frac{1}{2}1)$ | 4.02 | 0.24 | $(\frac{1}{2}\frac{1}{2}-1)$ | 5.46 | 0.038 |

The calculated values agree with the experiment in its order of magnitude as well as in the qualitative nature of its temperature dependence. Considering the approximate nature of our treatment, the agreement may be said to be satisfactory, though theoretical temperature dependence of T_1 is weaker than that observed. For the purpose of seeing the effect of anisotropy, we show by a dotted curve in Fig. 2 the values calculated without anisotropy. Comparing these two curves, one sees that the anisotropy makes the relaxation time longer at low temperatures, as one might expect.

The weaker temperature dependence of the theoretical values than the experimental ones may at least partly originate from the long wave approximation (2.28). Comparing

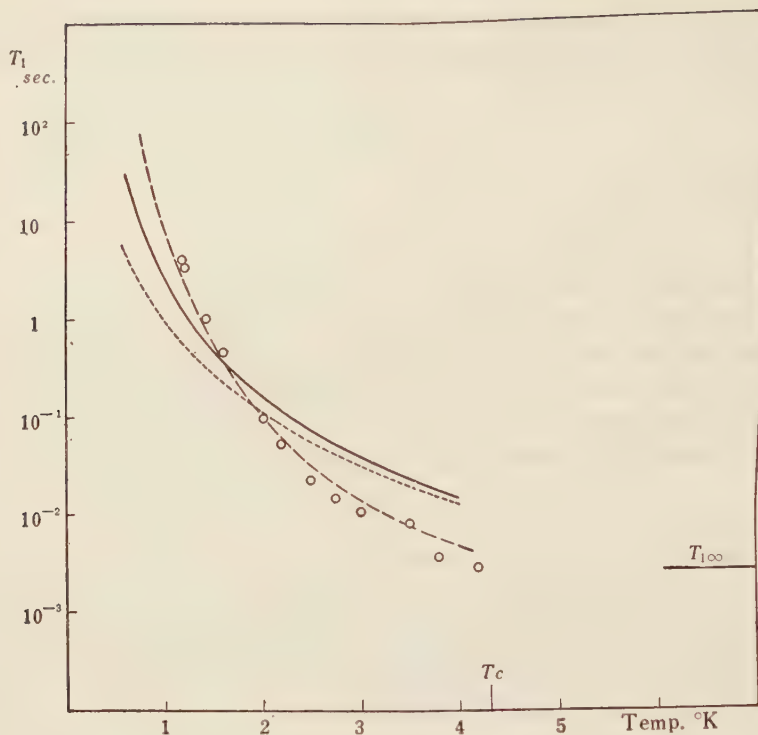


Fig. 2. T_1 for the protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$: the solid and the dotted curves are calculated by using long wave approximation with and without anisotropy for the energy spectra of the spin waves, the dashed curve is calculated assuming for the state density that given by (4.8) and the open circles are experimental points.

(2.28) with (2.6) we see that $N(\omega)$ calculated by the long wave approximation protrudes too much towards the higher frequency side. Also, the values of the specific heat and the sublattice magnetization calculated by this approximation do not well agree with experiments.

So we try to adjust $N(\omega)$ so as to make the theoretical curves of all these three quantities fit the corresponding experimental curves. It is known that the sublattice magnetization in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ deviates proportionally to T^1 from its value for absolute zero. Therefore we assume the state density to be given by

$$N(\omega) = \frac{5}{\omega_m} \omega^4. \quad (4.8)$$

T_1 then becomes

$$\frac{1}{T_1} = \frac{25}{2} \frac{\pi \gamma_e^2 \gamma_N^2 \hbar^3}{\kappa \theta} \sum_i r_i^{-6} F_i(\alpha \beta \gamma) \{ T^{*7} \mathfrak{S}_6(1/T^*) + T^{*9} \mathfrak{S}_8(1/T^*) \}, \quad (4.9)$$

where $\kappa\theta = \hbar\omega_m$, $T^* = T/\theta$ before, and $\mathfrak{N}_n(x)$ is defined by (2.34). Numerical values of $\mathfrak{N}_6(x)$ and $\mathfrak{N}_3(x)$ have been given by Ziman¹⁷⁾. If we take

$$\theta = 15^\circ\text{K}, \quad (4.10)$$

the calculated three quantities agree satisfactorily with the experiments, as shown in Fig. 2, 3 and 4. The value of θ given by (4.10) may be reasonable since $J_2S/\kappa = 13.5^\circ\text{K}$

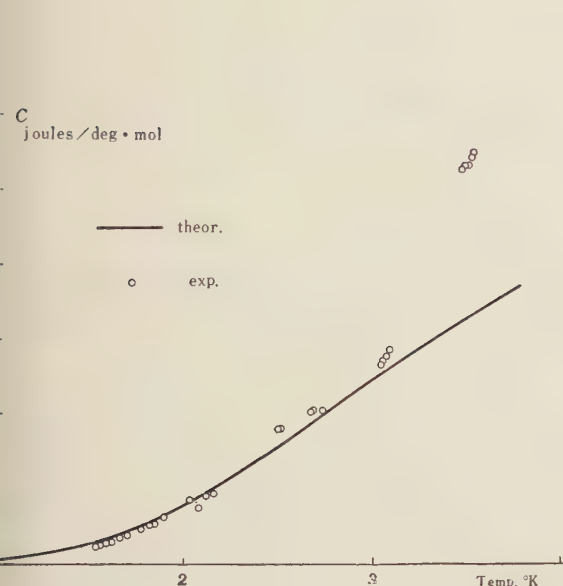


Fig. 3. Specific heat of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$.



Fig. 4. Sublattice magnetization of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$.

The relaxation time above the Curie point has not been observed. Its value at sufficiently high temperatures, $T_{1\infty}$, can be calculated by (3.16). ω_e may be calculated from (3.11) and (4.2) taking $z=2$, and is

$$\omega_e = 1.9 \times 10^{12}. \quad (4.11)$$

Taking the directional averages of $F_l(\alpha\beta\gamma)$ and $F'_l(\alpha\beta\gamma)$, one estimates

$$\sum_l r_l^{-6} [\langle F_l(\alpha\beta\gamma) \rangle + \langle F'_l(\alpha\beta\gamma) \rangle] = 4 \sum_l r_l^{-6} = 10^{16}. \quad (4.12)$$

Therefore we get

$$T_{1\infty} = 2.5 \times 10^{-3} \text{ sec.} \quad (4.13)$$

This is also shown in Fig. 2.

(II) MnF_2

This substance is of a rutile type structure. The Mn^{++} ions form a body centred

tetragonal lattice, the body centre sites and the corner sites being occupied by spins pointing to the + and -c directions, respectively, below the Curie point. Hence the theory of the previous sections is directly applicable to this substance.

The exchange integral is estimated as before from the observed perpendicular susceptibility¹⁸⁾ making use of (4.1). We have

$$Jz/\kappa=30.8^{\circ}\text{K}, \tag{4.14}$$

which corresponds, according to (2.29), to

$$\theta=150^{\circ}\text{K}. \tag{4.15}$$

Keffer¹⁹⁾ estimated the anisotropy field to $H_A=8800$ gauss, from which we have by (2.30)

$$T_{AE}=15^{\circ}\text{K}. \tag{4.16}$$

The lattice constants are²⁰⁾

$$a=4.8734, \quad b=3.3103, \quad u=0.310. \tag{4.17}$$

The F^- ions occupy four positions $(\pm u, \pm u, 0)$ and $(\frac{1}{2}\pm u, \frac{1}{2}\pm u, \frac{1}{2})$ in the unit cell. We consider an F^- ion at $(u, u, 0)$. Some numerical values of r_l^{-6} are given in Table III. Taking the directional average of $F_l(\alpha\beta\gamma)$ we get

$$\sum_l r_l^{-6} \langle F(\alpha\beta\gamma) \rangle = 5 \times 10^{46}. \tag{4.18}$$

Using these numerical values we calculate T_1 from (2.32) and (2.35). The results are shown in Fig. 5.

Next we shall turn to the paramagnetic region. Evaluating as before

$$\omega_e=1.7 \times 10^{12}, \tag{4.19}$$

and

$$\sum_l r_l^{-6} [\langle F_l(\alpha\beta\gamma) \rangle + \langle F_l'(\alpha\beta\gamma) \rangle] = 1.4 \times 10^{47}, \tag{4.20}$$

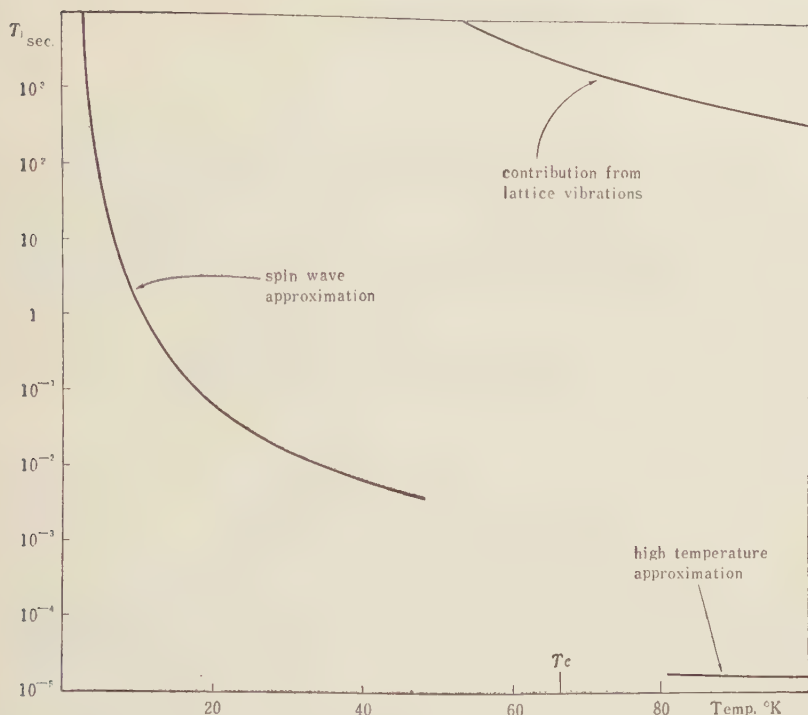
we get from (3.16)

Table III

| position of Mn^{++} | r_l (Å) | r_l^{-6} (cgs) | position of Mn^{++} | r_l (Å) | r_l^{-6} (cgs) |
|---------------------------------------|-----------|------------------------|------------------------------|-----------|------------------------|
| $(\frac{1}{2}\frac{1}{2}\frac{1}{2})$ | 2.101 | 11.32×10^{45} | (0 1 0) | 3.686 | 0.398×10^{45} |
| (0 0 0) | 2.137 | 10.51 | (00±1) | 3.940 | 0.268 |
| (1 0 0) | 3.686 | 0.398 | | | |

$$T_{1\infty}=1.6 \times 10^{-5} \text{ sec}. \tag{4.21}$$

The temperature dependence in this approximation is weak as seen in Fig. 5 and also in Fig. 1.

Fig. 5. Theoretical values of T_1 for F in MnF_2

We next concern ourselves with the effect of the lattice vibrations. The Curie point of MnF_2 is 66.5°K which is not very low compared with its Debye temperature θ_D . Stout and Catalano²¹⁾ have recently separated in an ingenious way the lattice and electronic part of the specific heat of the iron group fluorides; the corresponding entropies are of comparable order of magnitude below the Curie point. Hence it is necessary to estimate the contribution from the lattice vibrations to the nuclear relaxation. This subject has been studied by Waller.²²⁾ He took a simple cubic lattice of paramagnetic ions and calculated the thermal transition probability due to the dipole interactions among the spins. With slight modifications we can utilize his formula for estimating the order of magnitude of T_1 of our case. We have

$$(1/T_1)_{\text{lattice}} = 2 \sum_i \frac{245}{192\pi^5} \frac{\mu_e^2 \gamma_N^2 \hbar^2}{\rho^2 v^{10} r_i^6} \int_0^{\omega_{\max}} \frac{e^{\hbar\omega/kT} \omega^6}{(e^{\hbar\omega/kT} - 1)^2} d\omega, \quad (4.22)$$

where ρ denotes the density, v the sound velocity, μ_e the dipole moment of a Mn^{++} ion and ω_{\max} the maximum frequency of the lattice waves, this being connected with the Debye temperature by

$$\hbar\omega_{\max} = \kappa\theta_D.$$

Defining reduced temperature

$$T_D^* = T/\theta_D, \quad (4.23)$$

we get

$$(1/T_1)_{lattice} = \frac{490}{192\pi^5} \frac{\mu_e^2 \gamma_N^2 \hbar^2}{\mu^2 v^{10}} \sum_i \gamma_i^{-6} (\kappa \theta_D / \hbar)^7 T_D^{*7} \zeta_6(1/T_D^*). \quad (4.24)$$

In this expression μ_e should not always be considered as the true moment. The exchange interactions make spins fluctuate with time ($\omega \sim 10^{12}$) so that the averaged moment should be effective in processes concerned with low frequency phonons ($\omega < \omega_e$), while the true moment is to be effective in processes concerned with high frequency phonons ($\omega \gg \omega_e$). For simplicity, however, we shall here take the true moment for all the modes of lattice waves; by doing so, we obtain the lower limit of T_1 . Estimating

$$\rho = 4, v = 4 \times 10^5, \theta = 450, \quad (4.25)$$

we get

$$(1/T_1)_{lattice} = 0.6 T_D^{*7} \zeta_6(1/T_D^*). \quad (4.26)$$

Numerically evaluating this quantity, we find it to be smaller than the contribution from the exchange interactions by a factor of $10^{-5} \sim 10^{-7}$, thus being quite negligible.

Our final result is shown in Fig. 5. In view of the success of our theory in the case of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, Fig. 5 must be reliable at least as regards the order of magnitude. We see that T_1 is always longer than 10^{-5} sec. and becomes rapidly longer as temperature is lowered. This, however, contradicts with the experiment reported by Bloembergen and Poulis⁽³⁾; they used a rf field of 30.5 Mc/sec and the external fields covering 0 ~ 10000 gauss, but observed no resonance at temperatures between 300 K and 1.5° K. They attributed the absence of the resonance to a very short relaxation time ($< 10^{-6}$ sec), which, however, is unreasonably too short compared with our prediction.

As a possible origin of this discrepancy, a hyperfine interaction of magnetic electrons with fluorine nuclei may be suggested. The crystal of MnF_2 will not be perfectly ionic but will include a certain amount of covalency. The electrons on an F^- ion will occasionally move to one of the neighboring Mn^{++} ions whose spin points upwards or downwards according as it belongs to the plus or minus spin-sublattice. The theory of super-exchange tells that because of the exclusion principle, an electron with a given spin can be transferred only to those Mn^{++} ions whose spins are antiparallel to that electron spin. Moreover, the transfer rate of electrons from a given F^- to the only one nearest Mn^{++} ion should be different from that to the two next nearest Mn^{++} ions of the opposite spin. Accordingly a net electron spin moment should be left on each F^- ion, which interacts strongly with the fluorine nucleus. It is possible that this net moment makes the fluorine nuclear resonance frequency fall outside the range covered by Bloembergen and Poulis.

The hyperfine coupling constants of the 2s and 2p electrons of the free F^- ion are estimated as follows:

$$\begin{aligned} \frac{8}{3} \pi 2 \mu_B \gamma_N \hbar |\varphi_{2s}(0)|^2 &= 1.57 \text{ cm}^{-1}, \\ \frac{2}{3} 2 \mu_B \gamma_N \hbar \langle 1/r^3 \rangle_{2p} &= 0.044 \text{ cm}^{-1}, \end{aligned} \quad (4 \cdot 27)$$

where $\varphi_{2s}(0)$ denotes the value at the nucleus of the normalized $2s$ wave function whose numerical value can be obtained from the Hartree wave function²³⁾, $\langle 1/r^3 \rangle_{2p}$ the expectation value of r^{-3} with respect to the $2p$ function whose numerical value can be taken from Barnes-Smith's table.²¹⁾ We take a $2p$ orbital which extends perpendicularly to the c -axis. Correspondingly to (4·27), the magnetic field produced by a $2s$ or $2p$ electron at the nucleus is

$$\begin{aligned} H(2s) &= 6.1 \times 10^6 \text{ gauss}, \\ H(2p) &= 1.8 \times 10^5 \text{ gauss}. \end{aligned} \quad (4 \cdot 28)$$

When the external magnetic field is below 10000 gauss and the internal field due to the electron spins is above 17000 gauss, corresponding to a net $2s$ electron spin moment of larger than 0.3 percent or a net $2p$ spin moment of larger than 10 percent on the F^- nucleus, the nuclear resonance frequency is higher than 30.5 Mc/sec and falls outside the range of frequencies covered by Bloembergen-Poullis' experiment.

The net moments on F^- ions make T_1 shorter. It may be assumed that each of them fluctuates synchronously with the moment on the nearest manganese ion. The magnetic field produced at the fluorine nucleus by the hyperfine interaction is 20~60 times stronger than the dipolar field if there are unpaired $2s$ electrons of 0.3~1 percent; therefore T_1 should be shorter than the values given in Fig. 5 by a factor of $2 \times 10^{-3} \sim 3 \times 10^{-4}$. If this is actually the case, T_1 is too short above and near the Curie temperature to make the resonance observable, while at very low temperatures T_1 is still long enough to permit the observation. The amount of greater than 0.3 percent of unpaired $2s$ spins is not unreasonable.*

In conclusion the writer wishes to express his cordial thanks to Prof. T. Nagamiya and Dr. K. Yosida for their kind interest and valuable discussions.

Appendix

We wish to derive the formula (3·1) for T_1 by using perturbation theory. Following the notation used in the text, the unperturbed Hamiltonian is

$$\mathfrak{H}_0 = \mathfrak{H}_N + \mathfrak{H}_e, \quad (A \cdot 1)$$

with

$$\mathfrak{H}_N = -\hbar \omega_0 I_z, \quad (A \cdot 2)$$

*) Recently Tinkham reported an experimental evidence of such hyperfine interactions in the paramagnetic resonance of MnF_2 , FeF_2 and CoF_2 diluted with ZnF_2 . He observed a super-hyperfine structure of the resonance line which could be interpreted as arising from the interactions of the magnetic electrons with the nuclei of the surrounding F^- ions. [Phys. Rev. **100** (1955), 1792.]

ω_0 being the nuclear Larmor frequency. The perturbation is

$$\begin{aligned}
 \mathcal{H}' &= \gamma_e \gamma_N \hbar^2 \sum_l r_l^{-3} \mathbf{I} \cdot \left(\mathbf{1} - 3 \frac{\mathbf{r}_l \mathbf{r}_l}{r_l^2} \right) \cdot \partial \mathbf{S}_l \\
 &= \gamma_e \gamma_N \hbar^2 \sum_l r_l^{-3} [I_x \{ \alpha \partial S_{lx} + \beta \partial S_{ly} + \gamma \partial S_{lz} - 3(\alpha_l \alpha + \beta_l \beta + \gamma_l \gamma) \\
 &\quad \times (\alpha_l \partial S_{lx} + \beta_l \partial S_{ly} + \gamma_l \partial S_{lz}) \} \\
 &\quad + I^+ \{ \alpha^- \partial S_{lx} + \beta^- \partial S_{ly} + \gamma^- \partial S_{lz} - 3(\alpha^- \alpha_l + \beta^- \beta_l + \gamma^- \gamma_l) \\
 &\quad \times (\alpha_l \partial S_{lx} + \beta_l \partial S_{ly} + \gamma_l \partial S_{lz}) \} \\
 &\quad + I^- \{ \alpha^+ \partial S_{lx} + \beta^+ \partial S_{ly} + \gamma^+ \partial S_{lz} - 3(\alpha^+ \alpha_l + \beta^+ \beta_l + \gamma^+ \gamma_l) \\
 &\quad \times (\alpha_l \partial S_{lx} + \beta_l \partial S_{ly} + \gamma_l \partial S_{lz}) \}], \quad (\text{A} \cdot 3)
 \end{aligned}$$

with

$$I^\pm = I_x \pm iI_y, \quad \alpha^\pm = \frac{\alpha_1 \pm i\alpha_2}{2}, \text{ etc.}$$

Specifying the eigenstates of the nuclear spin by m and the electron spin system by ν , the transition probability corresponding to $(m', \nu) \rightarrow (m, \nu')$ is obtained as follows by the first order perturbation calculation:

$$W(m, \nu \rightarrow m', \nu') = \frac{1}{t} \cdot \frac{1}{\hbar^2} \int_0^t \int_0^{t'} dt' dt'' (m, \nu | \mathcal{H}'(t') | m', \nu') (m', \nu' | \mathcal{H}'(t'') | m, \nu) \quad (\text{A} \cdot 4)$$

where

$$\mathcal{H}'(t) = e^{it\mathcal{H}_0/\hbar} \mathcal{H}' e^{-it\mathcal{H}_0/\hbar}. \quad (\text{A} \cdot 5)$$

Inserting (A.3) in (A.4) and using the relation

$$e^{it\mathcal{H}_0/\hbar} I^\pm e^{-it\mathcal{H}_0/\hbar} = e^{\mp i\omega_0 t} I^\pm,$$

we get

$$\begin{aligned}
 W(m, \nu \rightarrow m+1, \nu') &= \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l \sum_{l'} r_l^{-3} r_{l'}^{-3} (I-m)(I+m+1) \\
 &\quad \times \frac{1}{t} \int_0^t \int_0^{t'} dt' dt'' e^{i\omega_0(t''-t')} (\nu | K_l^+(t'') | \nu') (\nu' | K_{l'}^-(t') | \nu), \quad (\text{A} \cdot 6)
 \end{aligned}$$

where

$$\begin{aligned}
 K_l^\pm(t) &= \alpha^\pm \partial S_{lx}(t) + \beta^\pm \partial S_{ly}(t) + \gamma^\pm \partial S_{lz}(t) \\
 &\quad - 3(\alpha^\pm \alpha_l + \beta^\pm \beta_l + \gamma^\pm \gamma_l) \{ \alpha_l \partial S_{lx}(t) + \beta_l \partial S_{ly}(t) + \gamma_l \partial S_{lz}(t) \}, \quad (\text{A} \cdot 7)
 \end{aligned}$$

$\partial S_l(t)$ being defined in § 3.

To obtain the thermal transition probability of a nuclear spin, we multiply (A.6) by a statistical weight of the initial state and sum up over the initial and final states. The result is

$$W(m \rightarrow m+1) = \gamma_e^2 \gamma_N^2 \hbar^2 (I-m)(I+m+1) \sum_l \sum_{l'} r_l^{-3} r_{l'}^{-3} \\ \times \frac{1}{t} \int_0^t \int_0^t dt' dt'' e^{i\omega_0(t''-t')} \langle K_l^+(t'') K_{l'}^-(t') \rangle, \quad (\text{A} \cdot 8)$$

where $\langle \rangle$ means the statistical average defined by (3.3). After simple manipulations, neglecting correlations between different spins, we get

$$W(m \rightarrow m+1) = \gamma_e^2 \gamma_N^2 \hbar^2 (I-m)(I+m+1) \sum_l r_l^{-6} \int_{-\infty}^{\infty} \langle K_l^+(t) K_l^- \rangle e^{i\omega_0 t} dt. \quad (\text{A} \cdot 9)$$

Similarly

$$W(m+1 \rightarrow m) = \gamma_e^2 \gamma_N^2 \hbar^2 (I-m)(I+m+1) \sum_l r_l^{-6} \int_{-\infty}^{\infty} \langle K_l^-(t) K_l^+ \rangle e^{-i\omega_0 t} dt. \quad (\text{A} \cdot 10)$$

Therefore T_1 becomes

$$\frac{1}{T_1} = \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l r_l^{-6} \int_{-\infty}^{\infty} \{ \langle K_l^+(t) K_l^- \rangle e^{i\omega_0 t} + \langle K_l^-(t) K_l^+ \rangle e^{-i\omega_0 t} \} dt. \quad (\text{A} \cdot 11)$$

From (A.7) we calculate

$$\langle K_l^{\pm}(t) K_l^{\mp} \rangle = \frac{1}{4} [F_l(\alpha\beta\gamma) \langle \partial S_{lx}(t) \partial S_{lx} \rangle + F_l'(\alpha\beta\gamma) \frac{1}{4} \{ \langle \partial S_l^+(t) \partial S_l^- \rangle + \langle \partial S_l^-(t) \partial S_l^+ \rangle \} \\ \pm F_l''(\alpha\beta\gamma) \frac{1}{4} \{ \langle \partial S_l^+(t) \partial S_l^- \rangle - \langle \partial S_l^-(t) \partial S_l^+ \rangle \}], \quad (\text{A} \cdot 12)$$

where $F_l(\alpha\beta\gamma)$ and $F_l'(\alpha\beta\gamma)$ are defined by (2.17) and (3.4) and $F_l''(\alpha\beta\gamma)$ is

$$F_l''(\alpha\beta\gamma) = -\gamma(1-\gamma_i^2) - \alpha\alpha_i\gamma_i + \beta\beta_i\gamma_i. \quad (\text{A} \cdot 13)$$

Inserting (A.12) into (A.11) and considering that $\sin \omega_0 t$ is negligible when $\langle \partial S_l(t) \partial S_l \rangle$ takes an appreciable value, we get

$$\frac{1}{T_1} = \frac{1}{2} \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l r_l^{-6} [F_l(\alpha\beta\gamma) \int_{-\infty}^{\infty} dt \cos \omega_0 t \langle \partial S_{lx}(t) \partial S_{lx} \rangle \\ + F_l'(\alpha\beta\gamma) \int_{-\infty}^{\infty} dt \cos \omega_0 t (1/4) (\langle \partial S_l^+(t) \partial S_l^- \rangle + \langle \partial S_l^-(t) \partial S_l^+ \rangle)]. \quad (\text{A} \cdot 14)$$

This is the required formula.

References

- 1) N. J. Poulis and G. E. G. Hardeman, *Physica* **18** (1952), 201.
- 2) N. J. Poulis and G. E. G. Hardeman, *Physica* **18** (1952), 315, 429; *ibid.* **19** (1953), 391.
- 3) N. Bloembergen and N. J. Poulis, *Physica* **16** (1950), 915.
- 4) G. E. G. Hardeman, *Conference de Physique des basses temperatures*, Paris, 1955.
- 5) S. A. Friedberg, *Physica* **18** (1952), 714.
- 6) R. Kubo, *Phys. Rev.* **87** (1952), 568.
- 7) T. Nakamura, *Prog. Theor. Phys.* **7** (1952), 539.
- 8) J. A. Eisele and F. Keffer, *Phys. Rev.* **96** (1954), 929.
- 9) N. Bloembergen, E. M. Purcell and R. V. Pound, *Phys. Rev.* **73** (1948), 679.
- 10) R. Kubo and K. Tomita, *J. Phys. Soc. Japan* **9** (1954), 888.
- 11) T. Nagamiya, K. Yosida and R. Kubo, *Advances in Physics* **4** (1955), 1.
- 12) J. Van den Handel, H. M. Gijsman and N. J. Poulis, *Physica* **18** (1952), 862.
- 13) D. Harker, *Z. Krist.* **93** (1936), 136.
- 14) J. Ubbink, J. A. Poulis, H. J. Gerritsen and C. J. Gorter, *Physica* **18** (1952), 361.
- 15) T. Nagamiya, *Prog. Theor. Phys.* **11** (1954), 309.
- 16) J. Itoh, R. Kusaka, Y. Yamagata, R. Kiriyaama and H. Ibamoto, *Physica* **19** (1953), 415.
- 17) J. M. Ziman, *Proc. Roy. Soc. A* **226** (1954), 436.
- 18) H. Bizette and B. Tsai, *Comptes Rendus* **238** (1954), 1575.
- 19) F. Keffer, *Phys. Rev.* **87** (1952), 608.
- 20) M. Griffel and J. W. Stout, *J. Amer. Chem. Soc.* **72** (1950), 4351.
- 21) J. W. Stout and E. Catalano, *J. Chem. Phys.* **23** (1955), 2013.
- 22) I. Waller, *Z. Phys.* **79** (1932), 370.
- 23) D. R. Hartree, *Proc. Roy. Soc.* **151** (1935), 96.
- 24) R. G. Barnes and W. V. Smith, *Phys. Rev.* **93** (1954), 95.

A Theory of Metallic Ferro- and Antiferromagnetism on Zener's Model

Tadao KASUYA

Physical Institute, Nagoya University, Nagoya

(Received February 3, 1956)

The importance to the mechanism of ferromagnetism of exchange interaction between conduction electrons and unfilled inner shell electrons (called s-d interaction) has been pointed out by Zener. Especially for rare earth metals, this interaction seems to be the only mechanism which can cause ferro- and antiferromagnetism. However Zener's works are unsatisfactory because his model is phenomenological and moreover does not involve antiferromagnetism and spin wave mode.

Our paper considers this s-d interaction on a more rigorous basis. By a certain approximation, there appear long range exchange type interactions between d-electron spins and, in certain conditions both ferro- and antiferromagnetism appear. The excitations of spin wave modes are the same as those in the ordinary modes of the short range exchange force, viz, the energy of the spin wave excitations is proportional to q^2 for ferromagnetism and q for antiferromagnetism in the region of small wave vector q . The $T^{3/2}$ law for the temperature dependence of the magnetization of ferromagnetism is applicable up to very high temperatures, and this result is in good agreement with the results of experiments on metallic ferromagnetism.

§ 1. Introduction

Since Heisenberg,¹⁾ there have been many discussions on the origin of the ferro- and antiferromagnetism; however, no satisfactory theory has so far been offered. The simplest theory is that of Heisenberg using atomic wave functions, a standpoint which may be suitable for non metallic substances. Even in such a case, however, there exist certain ambiguities, as was pointed out by Slater.²⁾ It is generally accepted that Heisenberg's model can not be adopted for metallic ferromagnetism where the situation is more difficult. It is conceivable that there are two different standpoints for the origins of ferro- and antiferromagnetism. One regards the exchange interaction between the inner shell electrons as very important to the origin of ferro- and antiferromagnetism, and considers the role of conduction electron to be negligible. The other regards the exchange interaction between the conduction electron and the inner shell electron as essential. The former interaction may be essential for transition metals, e. g., Ni, Fe, Co, Cr, Mn, etc., but for rare earth metals, this interaction almost completely vanishes and the latter interaction seems to increase in importance. Even for transition metals, the latter interaction exists and plays an important role in many phenomena; for example, in the relaxation process of microwave resonance absorption;^{3,4)} in the anomalous electrical resistance of transition metals,⁵⁾ and in the temperature dependence of the anisotropy energy of Ni,⁶⁾ as was previously discussed by the author. This interaction (s-d interaction) was first discussed by Zener.⁷⁾ However his treatment is only phenomenological and on many points unsatisfactory, because his

theory does not involve antiferromagnetism or spin wave excitation. In this paper we consider this interaction on a rigorous basis, and obtain some interesting results.

§ 2. Basic Hamiltonian

The original Hamiltonian is written as follows,

$$H = \sum_i p_i^2/2m + \sum_n \sum_i v_0(|\mathbf{r}_i - \mathbf{R}_n|) + \sum_{i>j} \sum_j e^2/r_{ij}, \quad (1)$$

where the first term is the kinetic energy of the electrons, the second the interaction between the electrons and the ions, and the third the Coulomb interaction between the electrons. Here we consider that the positions of ions are fixed in their equilibrium positions. Now we treat Eq. (1) by the method of the second quantization and expand the quantized wave functions as follows;

$$\begin{aligned} \psi(\mathbf{r}) &= \sum_t \sum_v a_{tv} \varphi_{tv}(\mathbf{r}) \\ \psi^*(\mathbf{r}) &= \sum_t \sum_v a_{tv}^* \varphi_{tv}^*(\mathbf{r}), \end{aligned} \quad (2)$$

where t and v represent the character of orbital state and the direction of spin respectively and $\varphi_{tv}(\mathbf{r})$ satisfies the following equation,

$$\{p^2/2m + v(\mathbf{r})\} \varphi_{tv}(\mathbf{r}) = \epsilon_t \varphi_{tv}(\mathbf{r}), \quad (3)$$

where the potential $v(\mathbf{r})$ is one to be determined later.

By using expansion (2), Hamiltonian (1) becomes

$$\begin{aligned} H = & \sum_t \sum_v \epsilon_t a_{tv}^* a_{tv} - \sum_t \sum_{t'} \sum_v a_{tv}^* \langle t|v-v_0|t' \rangle a_{t'v} \\ & + \frac{1}{2} \sum_{t_1} \sum_{t_2} \sum_{t_3} \sum_{t_4} \sum_v \sum_\mu a_{t_1v}^* a_{t_2\mu}^* \langle t_1, t_2 | e^2/r | t_3, t_4 \rangle a_{t_3\mu} a_{t_4v}. \end{aligned} \quad (4)$$

Now $v(\mathbf{r})$ is determined self-consistently as follows

$$v(\mathbf{r}) - v_0(\mathbf{r}) = \sum_t \sum_\mu \langle t; \mathbf{r}' | e^2/|\mathbf{r} - \mathbf{r}'| | t; \mathbf{r}' \rangle a_{t\mu}^* a_{t\mu}. \quad (5)$$

Among the remaining terms of the third term of (4), we neglect the interaction between the conduction electrons because this interaction gives the correlation and the weak exchange interaction, and is negligible compared with the exchange interaction between the conduction and the unfilled inner shell electrons. Among the remaining interactions between the unfilled shell electrons, the intra-atomic interaction gives a strong Hund coupling. On the other hand, the inter-atomic interaction is complicated for transition metals, while for rare earth metals it is negligibly small. We omit the term "inter-atomic" in this paper. The remaining interactions between the conduction electrons and the unfilled shell electrons are composed of two terms. The first term represents the transition between the conduction and the unfilled shell electrons, and this is of a higher order than the second term. We neglect the first term in this paper, but as this interaction seems to play some roles in the phenomena of rare earth metals, we will calculate it in a later papers. The

second term does not contain the transition between conduction and unfilled shell electrons and is written as follows

$$\begin{aligned} & \sum_{T_1 \neq T_2} \sum_{t_1 \neq t_2} \sum_{\nu} \sum_{\mu} A_{T_1 \nu}^* a_{t_1 \nu}^* \langle T_1, t_1 | e^2/r | t_2, T_2 \rangle a_{t_2 \mu} A_{T_2 \nu} \\ & + \sum_{T_1} \sum_{T_2} \sum_{t_1} \sum_{t_2} \sum_{\nu} \sum_{\mu} A_{T_1 \mu}^* a_{t_1 \mu}^* \langle T_1, t_1 | e^2/r | T_2, t_2 \rangle A_{T_2 \mu} a_{t_2 \nu} \end{aligned} \quad (6)$$

where large letters mean the operators of the unfilled shell electrons and small letters those of the conduction electrons.

The first term in (6) represents the correlation energy and does not depend on the direction of the spins; therefore we neglect this term. The second term represents the generalized exchange interaction between the conduction and the unfilled shell electrons, and depends on the direction of the spins. For rare earth metals except Gd, there remains an orbital moment and $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is a good quantum number; hence the situation is quite complicated. With this situation we shall treat in another paper. Here as our first assumption we consider only where the orbital moment vanishes. This condition is fulfilled in Gd and transition metals. Next, we assume that the wave functions of the conduction electrons are well approximated by the Bloch type wave functions and the wave functions of the unfilled shell electron are approximated by atomic wave functions neglecting the overlapping between the different atomic wave functions. Then for the creation and annihilation operators of unfilled shell electrons we can substitute the spin operators \mathbf{S}_n , and we get the following Hamiltonian

$$\begin{aligned} H_S = & \sum_i \sum_{\nu} \varepsilon_i a_{i\nu}^* a_{i\nu} - N^{-1} \sum_i \sum_{i'} \sum_n J(t, t') \exp [i(\mathbf{k}_i - \mathbf{k}_{i'}) \mathbf{R}_n] \\ & \times \{ (a_{i+}^* a_{i'+} - a_{i-}^* a_{i'-}) S_n^z + a_{i+}^* a_{i'-} S_n^- + a_{i-}^* a_{i'+} S_n^+ \}, \end{aligned} \quad (7)$$

where N is the number of magnetic lattice points in a unit volume, and

$$J(t, t') = N \int \int d\mathbf{r}_1 d\mathbf{r}_2 \varphi_{in}^*(r_1) \varphi_t^*(r_2) e^2/r_{12} \cdot \varphi_{in}(r_2) \varphi_{t'}(r_1) \exp [i(\mathbf{k}_{t'} - \mathbf{k}_t) \mathbf{R}_n] \quad (8)$$

does not depend on the lattice position \mathbf{R}_n . When an external field H_z is applied, we must add to (7) the Zeeman energy

$$H_{\text{Zeeman}} = -\beta H_z \{ 2 \sum_i (a_{i+}^* a_{i+} - a_{i-}^* a_{i-}) + g S_i^z \}. \quad (9)$$

A similar method was used first by Y. Hasegawa, which is based on atomic functions.

For simplicity, we assume in this paper that the energy spectrum of the conduction electron is free electron like, differing only in the effective mass m . Then the quantum number t in (7) is equal to the wave vector \mathbf{k} . Next, as $J(\mathbf{k}, \mathbf{k}')$ depends sensitively on $|\mathbf{k} - \mathbf{k}'|$ but on \mathbf{k} or \mathbf{k}' not sensitively, we assume that $J(\mathbf{k}, \mathbf{k}')$ depends only on $|\mathbf{k} - \mathbf{k}'|$. It is difficult to obtain the actual functional form of $J(q)$, but, it is clear that roughly it behaves as follows; as q increases from zero, $J(q)$ decreases slowly, and when q approaches the principal vector of the reciprocal lattice \mathbf{K} , $J(q)$ decreases rapidly to become very small. From this simplification, Eq. (7) can be written also in a different form. In-

roducing the well-known spin-wave operators

$$\begin{aligned} S_q &= N^{-1/2} \sum_n S_n \exp(iqR_n), \\ \sigma_q &= N^{-1/2} \sum_i \sigma_i \exp(iq\mathbf{r}_i), \end{aligned} \quad (10)$$

where \mathbf{r}_i and σ_i are the coordinate and the spin operator of i -th electron, respectively, we have

$$H_S = \sum_k \sum_v \epsilon_k a_{kv}^* a_{kv} - \sum_q J(q) S_q \sigma_{-q}, \quad (11)$$

or in the ordinary coordinates

$$H_S = \sum_i p_i^2 / 2m - \sum_n \sum_n J(|\mathbf{r}_i - \mathbf{R}_n|) S_n \sigma_i, \quad (12)$$

where

$$J(|\mathbf{r}_i - \mathbf{R}_n|) = N^{-1} \sum_q J(q) \exp[i(\mathbf{r}_i - \mathbf{R}_n) \cdot \mathbf{q}]. \quad (13)$$

The physical meanings of Eqs. (7), (11), (12) are quite distinct, namely, the Coulomb interaction between the conduction and the unfilled shell electron is periodic and does not scatter the conduction electron, but when the spin directions of unfilled shell electrons are disturbed, the exchange interactions are not periodic and do scatter the conduction electrons. The first order process gives rise to the resistivity, which effect has been calculated in another paper. The second order process gives rise to the effective spin-spin interaction between the unfilled shell electrons, which interaction we calculate in this work.

§ 3. The second order perturbation and the effective Hamiltonian

The Hamiltonian is

$$\begin{aligned} H_S &= \sum_k \hbar^2 k^2 / 2m (a_{k+}^* a_{k+} + a_{k-}^* a_{k-}) - \beta H_z (g S_z + \sigma_z) \\ &\quad - N^{-1} \sum_k \sum_{k'} \sum_n J(|\mathbf{k} - \mathbf{k}'|) \exp[i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{R}_n] \\ &\quad \times \{ (a_{k+}^* a_{k'+} - a_{k-}^* a_{k'-}) S_n^z + a_{k+}^* a_{k'-} S_n^- + a_{k'-}^* a_{k+} S_n^+ \}, \end{aligned} \quad (1)$$

where

$$\begin{aligned} S_z &= \sum_n S_n^z, \\ \sigma_z &= \sum_i \sigma_i^z = \sum_k (a_{k+}^* a_{k+} - a_{k-}^* a_{k-}). \end{aligned} \quad (2)$$

The energy of the first order perturbation is

$$H^{(1)} = -N^{-1} J(0) \sigma_z S_z. \quad (3)$$

The energy of the second order perturbation is obtained as follows

$$\begin{aligned}
 H^{(2)} = & -N^{-2} \left[\sum_{\mathbf{k}}' \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \right. \\
 & \times \{f^+(k')(1-f^+(k)) + f^-(k')(1-f^+(k)) S_n^z S_m^z \\
 & + \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^-(k')(1-f^+(k)) S_m^+ S_n^- + f^+(k')(1-f^-(k)) S_m^- S_n^+\} \left. \right] \quad (4)
 \end{aligned}$$

or, exchanging the \mathbf{k} and \mathbf{k}' , and rearranging

$$\begin{aligned}
 H^{(2)} = & -N^{-2} \sum_{\mathbf{k}}' \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^+(k') + f^-(k')\} S_n^z S_m^z \\
 & -N^{-2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^-(k') S_m^+ S_n^- + f^+(k') S_m^- S_n^+\} \\
 & + \sum_{\mathbf{k}} \sum_{\mathbf{k}'} J^2(|\mathbf{k}-\mathbf{k}'|) / (\hbar^2/2m) (k^2-k'^2) \{f^-(k')f^+(k) - f^+(k')f^-(k)\} S_z \quad (5)
 \end{aligned}$$

where $f^+(k)$ and $f^-(k)$ are the Fermi distribution function of plus and minus spin electrons, respectively, and prime in \sum' means that the summation does not include terms where $\mathbf{k}=\mathbf{k}'$, such terms having been already included in $H^{(1)}$. This calculation is similar to the nuclear $\mathbf{I}\mathbf{I}$ coupling in metals.⁸⁾ $H^{(2)}$ is now separable into two parts. One includes $\mathbf{R}_n=\mathbf{R}_m$ and is written in the form $\sigma_z S_z$ as follows

$$\begin{aligned}
 H_{SS}^{(2)} = & -N^{-2} \sum_{\mathbf{k}}' \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) / (\hbar^2/2m) (k^2-k'^2) \{f^+(k') + f^-(k')\} S_n^{z2} \\
 & -N^{-2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_n \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) / (\hbar^2/2m) (k^2-k'^2) \{f^+(k') + f^-(k')\} (S_n^{z2} + S_n^{y2}) \\
 & +N^{-2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} J^2(|\mathbf{k}-\mathbf{k}'|) / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^+(k')(1-f^-(k)) - f^-(k')(1-f^+(k))\} S_z. \quad (6)
 \end{aligned}$$

The last term in (6) is of a higher order when compared with $H^{(1)}$ or the remaining terms of $H^{(2)}$, because the expansion parameter of this perturbation is $2mJ(0)/\hbar^2 k_f^2$ where k_f is the wave vector of the Fermi surface or σ_z/N .

The remaining terms of $H^{(2)}$ are written as a sum of the spin-spin interactions between the different lattice points :

$$\begin{aligned}
 H_{SS}^{(2)} = & -N^{-2} \sum_{\mathbf{k}}' \sum_{\mathbf{k}'} \sum_{n \neq m} \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^+(k') + f^-(k')\} S_n^z S_m^z \\
 & -N^{-2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_{n \neq m} \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2) \\
 & \times \{f^+(k') + f^-(k')\} (S_n^x S_m^x + S_n^y S_m^y) \\
 & -N^{-2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \sum_{n \neq m} \sum_m J^2(|\mathbf{k}-\mathbf{k}'|) \exp[i(\mathbf{k}-\mathbf{k}')\mathbf{R}_{nm}] / (\hbar^2/2m) (k^2-k'^2)
 \end{aligned}$$

$$\times \{f^+(k') - f^-(k')\} (S_m^- S_n^+ - S_m^+ S_n^-), \quad (7)$$

where the last term is the higher order term and should be omitted in our approximation.

In the first and the second terms of (6) and (7) we can replace $\{f^+(k) + f^-(k)\}/2$ with the average distribution function $f(k)$ in our approximation and performing the summation over k , $H^{(2)}$ becomes

$$H^{(2)} = - \sum_n \sum_m J'(R_{nm}) S_n^z S_m^z - \sum_n \sum_m J(R_{nm}) (S_n^x S_m^x + S_n^y S_m^y), \quad (8)$$

where

$$J'(R_{nm}) = N^{-1} \sum_q J(q) \exp(iqR_{nm}), \quad (9)$$

$$J(R_{nm}) = N^{-1} \sum_q J(q) \exp(iqR_{nm}), \quad (10)$$

$$J(q) = (1/4a) J^2(q) f(q), \quad (11)$$

$$f(q) = 1 + (4k_f^2 - q^2)/4k_f q \cdot \ln(2k_f + q)/(2k_f - q), \quad (12)$$

$$a = 1/6 \cdot \hbar^2/m \cdot q_D^3/k_f, \quad (13)$$

$$J(0) = \lim_{q \rightarrow 0} J(q) = J^2(0)/2a, \quad (14)$$

and \sum' in (9) means the summation over all q , omitting $q=0$ in the terms $S_n^z S_m^z$, q_D is Debye's cut-off wave vector.

The spin dependent Hamiltonian is now

$$H_S = \sum_k \hbar^2 k^2 / 2m \cdot (a_{k+}^* a_{k+} + a_{k-}^* a_{k-}) - \beta H_z (y S_z + \sigma_z) \\ - N^{-1} J(0) \sigma_z S_z - \sum_n \sum_m J'(R_{nm}) S_n^z S_m^z - \sum_n \sum_m J(R_{nm}) (S_n^x S_m^x + S_n^y S_m^y). \quad (15)$$

In our approximation, the first term of this can be written in σ_z , and σ_z determined by the relation $\partial H_S / \partial \sigma_z = 0$. Then, using (14), we have

$$H_S = -\beta H_z y (1 + J(0)/ya) S_z - N(\beta H_z)^2 / 2a - \sum_n \sum_m J(R_{nm}) S_n^x S_m^x, \quad (16)$$

$$\sigma_z = (\beta H_z N + J(0) S_z) / a. \quad (17)$$

The relations (16), (17) are fundamental equations in the following discussions. The spin-spin interaction is thus of an exchange type.

§ 4. The ground states

As it is very difficult to obtain the exact solution for the ground state of the Hamiltonian (16), we discuss here only two physically important cases, that is, ferromagnetic and antiferromagnetic ordering cases.

When the ground state is ferromagnetic, all the spins of the unfilled shell electrons are parallel. This state is the exact solution of the Hamiltonian (16) and the ground state energy is

$$E_f = -(\beta H_z)^2 / 2a \cdot N - (g + J(0)/a) \beta H_z NS - NS^2 \sum_{\mathbf{K}} J(\mathbf{K}), \quad (1)$$

where \mathbf{K} are the principal vectors of the reciprocal lattice, including $\mathbf{K}=0$.

Hereafter we do not consider the effects of the external field. The effects of the external field will be discussed separately.

For the antiferromagnetic ordering, we assume a simplified model where the crystal lattice is separable into two equivalent sublattices in each of which all the spins are parallel, the magnetization of both sublattices, however, being antiparallel. Contrary to ferromagnetism, this state is not an exact ground state. As it is difficult to obtain the exact ground state of antiferromagnetism, we calculate as usual by the method of spin-wave approximation and, in the next section, we obtain energy correction due to the quantum effect in this way. The zero order energy, namely rigid spin model, is easily obtained as follows

$$\begin{aligned} E_a^0 &= -NS^2 \left[\sum_{\mathbf{R}_n} J(\mathbf{R}_n) - \sum_{\mathbf{R}_\mu} J(\mathbf{R}_\mu) \right] \\ &= -NS^2 / 2 \left[\sum_{\mathbf{G}} J(\mathbf{G}) - \sum_{\mathbf{K}} J(\mathbf{K}) + \sum_{\mathbf{Q}} J(\mathbf{Q}) \right] \\ &= -NS^2 \sum_{\mathbf{Q}} J(\mathbf{Q}), \end{aligned} \quad (2)$$

where \mathbf{R}_n and \mathbf{R}_μ are respectively the distance between the lattice points on the same sublattice and different sublattices, \mathbf{G} the principal vector of the sublattice, and group $\{\mathbf{Q}\}$ is a part of the group $\{\mathbf{G}\}$ and equal to $\{\mathbf{G}\} - \{\mathbf{K}\}$, thus excluding $\mathbf{Q}=0$.

The difference between E_a^0 and E_f is

$$E_a^0 - E_f = NS^2 \left[\sum_{\mathbf{K}} J(\mathbf{K}) - \sum_{\mathbf{Q}} J(\mathbf{Q}) \right]. \quad (3)$$

The sign of (3) depends sensitively on the dependence of $J(q)$ on q . Ferromagnetic ordering could be established when $J(q)$ decreases rapidly with increasing q .

As it is difficult to calculate $J(q)$ exactly, we can mention in our approximation only that s-d interaction can establish both ferromagnetism and antiferromagnetism.

We must study the thermal stability of these ground states by examining the energy spectrum of the excited states. This is done in the next section.

§ 5. The excited states of ferromagnetism

In this section we calculate the energy spectrum of the excited states when the ground state is ferromagnetic.

In low temperatures, actually the spin wave method is a good approximation. Now we introduce the well-known spin-wave operator

$$\begin{aligned} S_n^+ &= (2S)^{1/2} (1 - a_n^* a_n / 2S)^{1/2} a_n, \\ S_n^- &= (2S)^{1/2} a_n^* (1 - a_n^* a_n / 2S)^{1/2}, \\ S - S_n^z &= a_n^* a_n, \end{aligned} \quad (1)$$

and

$$\begin{aligned} a_n &= N^{-1/2} \sum_q \exp(-i \mathbf{q} \mathbf{R}_n) a_q, \\ a_n^* &= N^{-1/2} \sum_q \exp(i \mathbf{q} \mathbf{R}_n) a_q^*, \end{aligned} \quad (2)$$

then expanding with respect to the parameter $1/2S$ up to the term of the order of $(1/2S)^0$,

$$\begin{aligned} H_s &= -NS^2 + 2S \sum_q \sum_K [J(K) - J(|\mathbf{K} + \mathbf{q}|)] a_q^* a_q \\ &+ (2N^{-1}) \sum_{q_1} \sum_{q_2} \sum_{q_3} \sum_{q_4} \sum_K [J(|\mathbf{K} + \mathbf{q}_1|) + J(|\mathbf{K} + \mathbf{q}_1|) - 2J(|\mathbf{K} + \mathbf{q}_1 + \mathbf{q}_4|)] \\ &\quad \times \delta(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_3 - \mathbf{q}_4 + \mathbf{K}) a_{q_1}^* a_{q_2}^* a_{q_3} a_{q_4}. \end{aligned} \quad (3)$$

The first term in (3) is E_f in § 3. The second term is the excitation energy at absolute temperature zero. The third term is the first order correction at finite temperatures. In this section we consider only the second term. We are interested in the low temperature region, or in other words, in the small value region of wave vector q , and hence we can expand (3) by q .

Introducing the following variables

$$\begin{aligned} \mathbf{x} &= \mathbf{q}/2k_f, \quad \mathbf{p} = \mathbf{K}/2k_f, \quad X = p^2, \\ \delta &= \{(\mathbf{K} + \mathbf{q})^2 - K^2\}/4k_f^2 = 2\mathbf{p} \cdot \mathbf{x} + x^2, \end{aligned} \quad (4)$$

the energy of the excited states of the spin wave of wave vector \mathbf{q} is

$$\begin{aligned} \varepsilon_q^0 &= 2S \sum_K [J(K) - J(|\mathbf{K} + \mathbf{q}|)] \\ &= -2S \sum_p \left\{ \delta \frac{d}{dX} + \frac{1}{2} \delta^2 \frac{d^2}{dX^2} + \frac{1}{6} \delta^3 \frac{d^3}{dX^3} + \dots \right\} J(X). \end{aligned} \quad (5)$$

Considering the crystal is to be cubic, and taking terms up to q^4 , we have

$$\varepsilon_q^0 = -2S \sum_p \left\{ x^2 \left(\frac{d}{dX} + \frac{2}{3} X \frac{d^2}{dX^2} \right) + \frac{1}{2} x^4 \left(\frac{d^2}{dX^2} + \frac{4}{3} X \frac{d^3}{dX^3} \right) \right\} J(X). \quad (6)$$

To perform further calculations, we neglect the derivative of $J(q)$. This neglect may be a fairly good approximation when $\mathbf{K} \rightarrow 0$, but not so good when $\mathbf{K} \neq 0$. Therefore this approximation may be justified when the contribution to ε_q^0 is mainly from the part $\mathbf{K} = 0$. Then

$$\varepsilon_q^0 = A_2^0 (q/q_D)^2 + A_4^0 (q/q_D)^4 \quad (7)$$

$$\begin{aligned} A_2^0 &= \frac{S}{8a} \left(\frac{q_D}{k_f} \right)^2 \sum_p J^2(X) \left(\frac{1}{6X^{1/2}} \right) \ln \frac{X'^2 + 1}{X^{1/2} - 1} - \frac{1}{3} \frac{1}{X - 1} \\ &= \frac{S}{12a} \left(\frac{q_D}{k_f} \right)^2 \left[\sum_{X \sim 1} J^2(X) \left(1 + \frac{2}{3} X + \dots \right) \right. \\ &\quad \left. - \frac{1}{3} \sum_{X \gg 1} J^2(X) \frac{1}{X^2} \left(1 + \frac{6}{5X} + \dots \right) \right] \end{aligned} \quad (8)$$

$$\begin{aligned}
 A_1^0 &= \frac{S}{64a} \left(\frac{q_D}{k_f} \right)^4 \sum_p J^2(X) \left(-\frac{1}{8} \frac{X+7}{X^{3/2}} \ln \frac{X^{1/2}+1}{X^{1/2}-1} - \frac{1}{4} \frac{1}{X^2} \frac{X-7}{X-1} + \frac{2}{3} \frac{1}{X} \frac{1}{(X-1)^2} \right) \\
 &= (S/240a) (q_D/k_f)^4 \left[\sum_{X < 1} J^2(X) (1+X+\dots) \right. \\
 &\quad \left. + 5/2 \sum_{X > 1} J^2(X) X^{-1} (1+2X^{-1}+\dots) \right]. \tag{9}
 \end{aligned}$$

The condition of thermal stability of ferromagnetism is $\mathcal{E}_q^0 > 0$, or $A_2^0 > 0$. If the part $\mathbf{K} \neq 0$ were to make a larger contribution than that of $\mathbf{K} = 0$, the condition of the thermal stability would be $X < 1$, or in other words, $q_D/k_f < 1$, i. e., that the valency of the conduction electron must be larger than 2. In thermal stability, the part $X > 1$ contributes only almost zero and we can neglect this part.

Considering only the first term of the expansion of part $X < 1$,

$$A_1^0/A_2^0 = (1/20) (q_D/k_f)^2. \tag{10}$$

For comparison, we give the result of the ordinary nearest neighbors interaction. In this case

$$A_1^0/A_2^0 = - (1/20) (q_D R_0)^2 \tag{11}$$

where R_0 is the distance between the nearest neighbouring spins.

For simplicity, we put R_0 as a diameter of the sphere having the same volume as the specific volume per magnetic lattice point. Then

$$(q_D R_0)^2 = 23. \tag{12}$$

Therefore (11) is considerably larger than (10), i. e., the $T^{3/2}$ law of temperature dependence of the magnetization may hold very well. We have calculated this phenomenon in the next section.

§ 6. Temperature dependence of the magnetization

Recently Schafroth⁽⁹⁾ calculated the temperature dependence of magnetization by the variational method based on Kubo's method.⁽¹⁰⁾ According to Schafroth the $T^{3/2}$ law does not hold except at very near zero.

Such an implausible result is owing to the variational method, because by this method the treatment of the diagonal and non-diagonal parts is quite at disparity and thus the character of the exchange interaction is disregarded. Therefore an effective anisotropic field appears and the excitation of the spin-waves is unnaturally depressed. This is the reason for such an implausible result. To obtain reasonable results we must recognize that the non-diagonal terms are important and that in the approximation of spin wave method we must use an expansion with respect to the parameter $1/2S$. A detailed discussion was given in a previous paper.⁽¹¹⁾

Here we calculate the departure of the temperature dependence of the magnetization from the $T^{3/2}$ law caused by the following three mechanisms, each up to the first order; (i) the existence of the terms proportional to q^4 in the excitation energy \mathcal{E}_q ; (ii) the

higher order terms in the expansion of the spin wave method, namely, the third term in eq. (3) of § 4; (iii) the restriction that the degree of freedom of motion of the spin system is finite.

Now we consider the third term in eq. (3) of § 4.

The excitation energy of the spin wave of wave vector q contributed by the third term in (3) of § 4 is

$$J_q = 2 \sum_{q'} n_{q'}/N \cdot \sum_K [J(|\mathbf{K} + \mathbf{q}|) + J(|\mathbf{K} + \mathbf{q}'|) - J(K) - J(|\mathbf{K} + \mathbf{q} + \mathbf{q}'|)]. \quad (1)$$

Expanding in q and using the same method as in § 4, we have

$$A_q = -\frac{10}{3} S \sum_{x'} \frac{n_{q'}}{NS} x'^2 x^2 \sum_p \left(\frac{d^2}{dX^2} + \frac{4}{3} X \frac{d^3}{dX^3} \right) J(X), \quad (2)$$

where

$$x' = q'/2k_f, \quad (3)$$

and the others have the same meaning as in § 4.

The form (2) is the same as the second term in (6), § 4, and written with A_4^0 as

$$A_q = \frac{10}{3} A_4^0 \left(\frac{q}{q_D} \right)^2 \sum_{q'} \left(\frac{q'}{q} \right)^2 \frac{n_{q'}}{NS}. \quad (4)$$

Adding this term, the excitation energy of the spin wave of wave vector q becomes

$$\epsilon_q = A_2 (q/q_D)^2 + A_4^0 (q/q_D)^4, \quad (5)$$

where

$$\begin{aligned} A_2 &= A_2^0 \left\{ 1 + \frac{10}{3} \frac{A_4^0}{A_2^0} \sum_{q'} \left(\frac{q'}{q_D} \right)^2 \frac{n_{q'}}{NS} \right\} \\ &= A_2^0 \left\{ 1 + \frac{1}{6} \sum_{q'} \left(\frac{q'}{q_D} \right)^2 \frac{n_{q'}}{NS} \right\} \end{aligned} \quad (6)$$

$$A_4^0 = (1/20) (q_D/k_f)^2 A_2^0 \quad (7)$$

$$A_2^0 = (S/12a) (q_D/k_f)^2 \sum_K J^2(K) \quad (8)$$

and the magnetization in finite temperature T is

$$NS - S_z = \sum_{q=0}^{q_D} [\exp(\epsilon_q/\kappa T) - 1]^{-1}. \quad (9)$$

Replacing summation by the integration, and confining us only to the first order for the deviation of each of the three causes, we get

$$NS - S_z = 3.5 \left(\frac{\kappa T}{A_2^0} \right)^{3/2} \left[1 - 1.9 \frac{\kappa T}{A_2^0} \frac{A_1^0}{A_2^0} + 1.5 \frac{A_2^0 - A_2}{A_2^0} \right]$$

$$-0.43 \left(\frac{A_2^0}{\kappa T} \right)^{1/2} \exp(-A_2^0/\kappa T) \Big]. \quad (10)$$

Putting the values of (6) and (7), we have

$$NS - S_z = 3.5 \left(\frac{\kappa T}{A_2^0} \right)^{3/2} \left[1 - 0.095 \left(\frac{q_D}{k_f} \right)^2 \frac{\kappa T}{A_2^0} - \frac{0.67}{S} \left(\frac{q_D}{k_f} \right)^2 \left(\frac{\kappa T}{A_2^0} \right)^{5/2} - 0.43 \left(\frac{A_2^0}{\kappa T} \right)^{1/2} \exp(-A_2^0/\kappa T) \right]. \quad (11)$$

For example, taking $\kappa T/A_2^0 = 1/2$, the bracket in (11) is

$$[1 - 0.05 (q_D/k_f)^2 - (0.21/S) (q_D/k_f)^2 - 0.08] \quad (12)$$

and the deviation is less than 15% for the usual metal. The result that the $T^{3/2}$ law holds very well is in good agreement with the experimental results for rare earth metals, especially for Gd.¹²⁾

Contrary to the above result, the $T^{3/2}$ law does not hold so well when we use the ordinary nearest neighbors interaction method. In that case

$$A_4^0/A_2^0 = -(1/20) (q_D R_0)^2, \quad (13)$$

$$A_2 = A_2^0 \{1 - (0.45/S) (q_D R_0)^2 (\kappa T/A_2^0)^{5/2}\}, \quad (14)$$

and, as mentioned in § 4, by approximating

$$(q_D R_0)^2 = 23, \quad (15)$$

the bracket in (10) becomes

$$\left[1 + 2.2 \left(\frac{\kappa T}{A_2^0} \right) + \frac{15}{S} \left(\frac{\kappa T}{A_2^0} \right)^{5/2} - 0.43 \left(\frac{A_2^0}{\kappa T} \right)^{1/2} \exp(-A_2^0/\kappa T) \right]. \quad (16)$$

The deviation is very much larger than that in (11) and the $T^{3/2}$ law holds only in low temperatures. The peculiar term which appeared in Schafroth's method does not appear in ours.

§ 7. The excited states of antiferromagnetism

For antiferromagnetism, we also introduce the spin wave operators a_q^* , a_q and b_q^* , b_q , which are the creation and annihilation operators, respectively, at A and B sites. Then taking terms up to the two spin wave process, in the expansion, we have

$$\begin{aligned} H_S = & -NS^2 \sum_Q J(Q) \\ & + 2S \sum_q \left[\sum_G \{J(G) - J(|G+q|)\} - \sum_Q J(Q) - \sum_K J(K) \right] (a_q^* a_q + b_q^* b_q) \\ & - 2S \sum_q \left[\sum_Q J(|Q+q|) - \sum_K J(|K+q|) \right] a_q b_{-q} + \text{c. c.} \end{aligned} \quad (1)$$

We easily diagonalize (1) by the well-known transformation, and (1) is written in

the new operators α_q^* , α_q , β_q^* and β_q , as

$$H_S = E_a + \sum_q \varepsilon_q (\alpha_q^* \alpha_q + \beta_q^* \beta_q), \quad (2)$$

where

$$E_a = -NS^2 \sum_Q J(Q) - 2S \sum_q \left[\sum_G \{J(G) - J(|G+q|)\} + \sum_Q J(Q) - \sum_K J(K) \right] - \sum_q \varepsilon_q \quad (3)$$

$$\varepsilon_q = 4S \left[\left\{ \sum_Q (J(Q) - J(|Q+q|)) \right\} \left\{ \sum_Q J(Q) - \sum_K J(K) + \sum_K (J(K) - J(|K+q|)) \right\} \right]^{1/2}. \quad (4)$$

To perform this transformation, it is required that (i) the coefficient of the second term of (1) be positive, and (ii) ε_q in (4) be real.

From § 3 and § 4, it is easily seen that

$$2S \sum_K \{J(K) - J(|K+q|)\} = A_2^0 (q/q_D)^2, \quad (5)$$

$$2S \sum_Q \{J(Q) - J(|Q+q|)\} = B_2^0 (q/q_D)^2, \quad (6)$$

$$2S \left\{ \sum_Q J(Q) - \sum_K J(K) \right\} = (2/NS) (E_f - E_a^0) \equiv \Delta E, \quad (7)$$

where B_2^0 is obtained from A_2^0 by replacing K with Q .

By using (5), (6) and (7)

$$E_a = E_a^0 - \sum_q \left[\{\Delta E + A_2^0 (q/q_D)^2\}^{1/2} - (B_2^0)^{1/2} q/q_D \right]^2, \quad (8)$$

$$\varepsilon_q = 2 \left[B_2^0 \{\Delta E + A_2^0 (q/q_D)^2\} \right]^{1/2} q/q_D, \quad (9)$$

and the restrictions (i) and (ii) become

$$\begin{aligned} \text{(I)} \quad & \Delta E > 0, \\ \text{(II)} \quad & A_2^0 > 0, \\ \text{(III)} \quad & B_2^0 > 0. \end{aligned} \quad (10)$$

(I) is fulfilled when antiferromagnetic state is more stable than ferromagnetic in the ground state energy. (II) is the same as in § 4. (III) is usually fulfilled when (II) is fulfilled or, strictly speaking, when the valency of the conduction electron is larger than unity. These three restrictions are the same as of thermal stability of antiferromagnetism.

§ 8. Conclusion

Here we summarize the above calculations.

(i) We assume that the conduction electron is nearly free electron like and that the unfilled inner shell electron is localized and the interaction between the unfilled shell electrons in the different lattice points is small.

(ii) We assume that the orbital moment of the unfilled shell electrons is quenched.

This assumption is true for transition metals because in these metals the crystalline field is very strong. In rare earth metals, however, l - s coupling is more important than the crystalline field and j is a good quantum number and hence the above assumption holds only for Gd. We will treat in another paper the case in which j is a well considered quantum number.

(iii) We assume the orbital configuration of the unfilled inner shell electrons to be fixed. This is related to (ii). The effects of the excited configuration seem to be small.

(iv) Under the above assumptions, the exchange interaction between the conduction and the unfilled inner shell electrons is written as a form of exchange type interaction as in Eqs. (7), (11) and (12) in § 2. The second order perturbation gives rise to the effective exchange-like interaction between the spins in the different lattice points. The character of this exchange-like interaction is a long range force and hence it is difficult to obtain the real ground state. The ordering of the ground state depends sensitively on the functional form of the generalized exchange integral $J(q)$.

(v) When the ground state is ferromagnetic, the excited state is written by spin wave mode, and the excited energy is proportional to the square of the wave vector. The $T^{3/2}$ law of temperature dependence of the magnetization holds very well, which is satisfactory to the experimental fact of Gd.

(vi) The excitation energy in the antiferromagnetic ordering is proportional to the absolute value of the wave vector.

(vii) For the thermal stability, it is required that the valence of the conduction electron be larger than 2.

The phenomena in fairly high temperatures and in the paramagnetic region will be treated in a following paper.

The author expresses his sincere appreciation to Prof. Ariyama for his encouragement. He also wishes to thank Prof. Matsubara for his many valuable discussions.

References

- 1) W. Heisenberg, *Zeits. f. Phys.* **49** (1928), 619.
- 2) J. C. Slater, *Phys. Rev.* **49** (1936), 537, 981; *ibid.* **52** (1937), 198; *ibid.* **82** (1951), 538
- 3) T. Kasuya, *Prog. Theor. Phys.* **12** (1954), 803.
- 4) E. Abrahams, *Phys. Rev.* **98** (1955), 387.
- 5) T. Kasuya, *Euseiron Kenkyu* (in Japanese) **92** (1956), 1.
- 6) T. Kasuya, *Euseiron Kenkyu* (in Japanese) **92** (1956), 7.
- 7) C. Zener, *Phys. Rev.* **81** (1951), 446; *ibid.* **82** (1951), 403; *ibid.* **83** (1951), 299; *ibid.* **85** (1951), 324.
- 8) M. A. Ruderman and C. Kittel, *Phys. Rev.* **96** (1954), 99.
- 9) M. R. Schafroth, *Proc. Phys. Soc.* **67** (1954), 33.
- 10) R. Kubo, *Rev. Mod. Phys.* **25** (1953), 344.
- 11) T. Kasuya, *Euseiron Kenkyu* (in Japanese) **92** (1956), 14.
- 12) L. F. Elliott, S. Legvold and F. H. Spedding, *Phys. Rev.* **91** (1953), 28.

Electrical Resistance of Ferromagnetic Metals

Tadao KASUYA

Physical Institute, Nagoya University, Nagoya

(Received February 13, 1956)

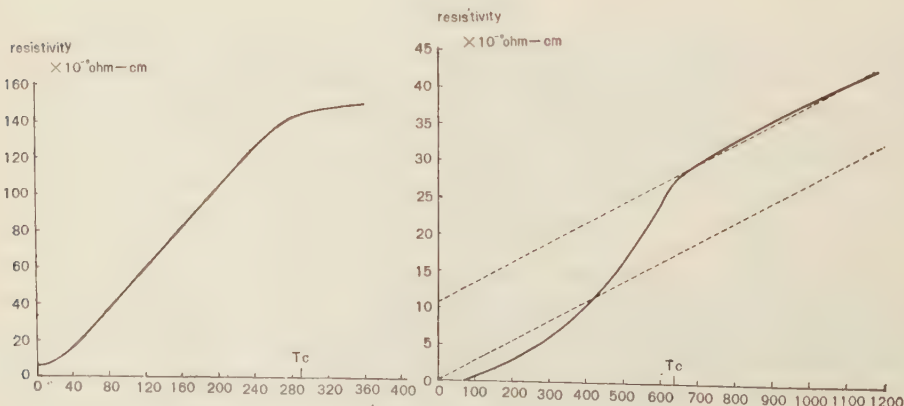
In ferromagnetic metals an anomalous electrical resistance is observed which is characteristic of the magnetization and is additive to the ordinary electrical resistance. The origin of this anomalous electrical resistance has been discussed by several authors, but all such attempts do not seem to have satisfactorily explained this phenomenon either quantitatively or qualitatively.

In this paper we calculate the anomalous electrical resistance from the standpoint of s-d interaction as developed by the author. It is our opinion that the anomalous electrical resistance occurs because the exchange energy between the conduction and the unfilled inner shell electrons depends on the relative direction of the spins of the electrons, and that this interaction is not periodic in finite temperatures.

In numerical values as well as in the temperature dependence, our results are in good agreement with experimental results.

§ 1. Introduction

In transition metals, Ni, Fe, etc., and also in rare earth metals, especially Gd, the electrical resistance depends on temperature as shown in Figs. 1¹ and 2²⁾. We can divide this resistance into two parts. The first is the normal resistance, proportional to the absolute temperature T in high temperatures. This temperature dependence suggests that this part of the resistance may be caused by electron-phonon interaction. The second is characteristic of ferromagnetism. The main feature of the temperature dependence of this



part is that when the temperature is higher the Curie temperature T_c , the resistance is almost constant, and when the temperature is lower than T_c , the resistance decreases to become zero at absolute temperature zero.

So far the origin of this anomalous electrical resistance has been discussed by Mott, etc.,³⁾ from the following standpoint. "In metallic ferromagnetics, one must consider two energy bands, that is, s-band and d-band. The effective mass of the d-band electron is very much larger than that of the s-band electron, and almost the entire electrical current is carried by s-electrons. Nevertheless, as the state density of the d-band is very large, the transition of electrons from s-band into d-band as caused by the electron-phonon interaction, is much more frequent than that from s-band into s-band. Thus the electrical resistance of transition metal is submitted to this s-d interband transition, and such a process depends on the magnetization of the d-band."

This process may actually exist and be important in the transition metals, but we think because of the following facts this should rather be included in the normal part of the electrical resistance, and that it does not correspond to the anomalous part in question. The reasons are as follows. (i) In rare earth metals the unfilled shell, that is the f-shell, does not seem to constitute a band, but behaves as if it were a free ion. Thus the above inter-band transition process does not occur. As is shown in Fig. 2, however, the anomalous resistance does occur in rare earth metals too. (Even in transition metals, it does not seem to be a good approximation to treat d-electron to form a band). (ii) The resistance caused by this process should be nearly proportional to the absolute temperature in high temperatures because this process occurs through the interaction with phonons. Such a temperature dependence, however, does not seem to explain the experimental facts of the anomalous electrical resistance, it rather corresponds to the normal one. (iii) Further, this process seems to be unable to explain satisfactorily the abrupt decrease of electrical resistance at the Curie point.

Our standpoint for this phenomenon is based on the s-d interaction described in detail in the former paper [I]¹⁾. We consider that s-electron (corresponding to s-and d-electrons in the rare earth metals) is nearly free and carries current, but the unfilled inner shell electron (d-electron in the transition metals and f-electron in the rare earth metals) does not carry current, being nearly localized. The Coulomb interaction between the conduction and the unfilled shell electrons has the same periodicity as that of the crystal, but the exchange interaction depends on the relative orientation of the spins of both electrons. Therefore at absolute temperature zero, all the spins of d-electrons (we mean by "d-electron" the unfilled shell electron) being in order, there is no resistance, while in finite temperatures this order is disturbed and the exchange interaction is no more periodic, and thus a resistance appears and increases with temperature. Above the Curie temperature, the direction of d-electron spin becomes perfectly at random, and resistance remains constant. This temperature dependence is exactly the same as that experimentally observed.

In the following sections we calculate qualitatively the electrical resistance caused by the above process.

§ 2. The basic equations

The basic Hamiltonian of the exchange interaction between the conduction and the unfilled shell electrons was discussed in detail in [I] and is written

$$H_{int} = -1/N \cdot \sum_{k_1} \sum_{k_2} \sum_{R_n} J(|\mathbf{k}_1 - \mathbf{k}_2|) \exp[i(\mathbf{k}_1 - \mathbf{k}_2) \mathbf{R}_n] \\ \times \{ (a_{k_1+}^* a_{k_2+} - a_{k_1-}^* a_{k_2-}) S_n^z + a_{k_1+}^* a_{k_2-} S_n^- + a_{k_1-}^* a_{k_2+} S_n^+ \} \quad (1)$$

where $a_{k\pm}^*$ and $a_{k\pm}$ are the creation and the annihilation operators of the state of wave vector \mathbf{k} of \pm spin, S_n is the spin operator of the unfilled shell electrons located at the lattice point \mathbf{R}_n , and $J(|\mathbf{k}_1 - \mathbf{k}_2|)$ is the generalized exchange energy

$$J(|\mathbf{k}_1 - \mathbf{k}_2|) = N \int d\mathbf{r}_1 d\mathbf{r}_2 \varphi_{k_1}^*(\mathbf{r}_1) \varphi_{k_2}(\mathbf{r}_2) \frac{e^2}{r_{12}} \varphi_{-k_2}(\mathbf{r}_2) \varphi_{-k_1}(\mathbf{r}_1) \exp[i(\mathbf{k}_1 - \mathbf{k}_2) \mathbf{R}_n]. \quad (2)$$

Here, $\varphi_{k\pm}(\mathbf{r})$ and $\varphi_{un}(\mathbf{r})$ are the wave functions of the conduction electron of wave vector \mathbf{k} and the unfilled shell electron at the point \mathbf{R}_n , and (2) reduces to the ordinary exchange energy when $\mathbf{k}_1 = \mathbf{k}_2$. As has already been mentioned, $J(|\mathbf{k}_1 - \mathbf{k}_2|)$ depends originally not only on $|\mathbf{k}_1 - \mathbf{k}_2|$ but also slightly on \mathbf{k}_1 or \mathbf{k}_2 where, however, the wave vector \mathbf{k} is always a vector on the Fermi surface, and hence we can completely neglect \mathbf{k}_1 dependence. Eq. (1) is not applicable when the orbital moment is not quenched.

In the following calculation, we use the approximation of the molecular field because this method is applicable in the whole temperature range and especially convenient to observe the behavior at the Curie temperature.

Above the Curie temperature, each spin S_n behaves perfectly independently and the conduction electrons are scattered at each lattice point. This situation somewhat resembles the mechanism of the electrical resistance of a perfectly disordered alloy. When the temperature is lower than the Curie temperature, an ordering appears and the situation becomes quite complicated. We must distinguish between the distribution functions of the conduction electrons of \pm spins, namely n_k^\pm . In the approximation that the Fermi energy is much larger than the exchange energy $J(0)$ and the temperature (in the unit of energy), we can consider that the Fermi energies of \pm spins, ξ_0^\pm are the same because of the equilibrium between the kinetic energy and the exchange energy, as was discussed in [I]. In the first term of (1), a scattering can be caused by the deviation of S_n^z from the average orientation σ , namely $(S_n^z - \sigma)$. From (1) we can easily see that in perfect ordering there is no scattering due to the first term. The second and the third terms represent the quantum effect and do not appear in alloys. When S_n^z changes by ± 1 because of these processes, the change in exchange energy can be written in the approximation of the molecular field as $\pm \alpha \sigma$, where α is related to the Curie temperature T_c as shown by

$$\alpha = (3/S(S+1)) \kappa T_c. \quad (3)$$

From the above consideration we can derive our Bloch equation. For simplicity, however, we calculate the resistance not by solving the Bloch equation, namely the equation

of the detailed balance, but by solving the equation of the balancing of the total wave vector \mathbf{K} .

The increase of \mathbf{K} due to an external field is

$$(d\mathbf{K}/dt)_{field} = n(e/\hbar) \mathbf{F}. \quad (4)$$

The increase of \mathbf{K} due to the scattering is

$$\begin{aligned} \left(\frac{d\mathbf{K}}{dt} \right)_{scatt} &= \frac{2\pi}{\hbar} N \sum_{\mathbf{k}_1} \sum_{\mathbf{k}_2} \sum_{S_z = -S}^S (\mathbf{k}_1 - \mathbf{k}_2) \\ &\times [|W_{k_1 k_2}|^2 \{ (1 - n_{k_1}^+) n_{k_2}^+ \partial(\varepsilon_{k_1}^+ - \varepsilon_{k_2}^+) + (1 - n_{k_1}^-) n_{k_2}^- \partial(\varepsilon_{k_1}^- - \varepsilon_{k_2}^-) \} (S_z - \sigma)^2 w_{sz} \\ &+ |V_{k_1 k_2}|^2 (1 - n_{k_1}^-) n_{k_2}^+ (S - S_z) (S + S_z + 1) w_{sz} \partial(\varepsilon_{k_1}^- - \varepsilon_{k_2}^+ - \alpha\sigma) \\ &- |V_{k_1 k_2}|^2 (1 - n_{k_2}^+) n_{k_1}^- (S - S_z) (S + S_z + 1) w_{sz+1} \partial(\varepsilon_{k_2}^+ - \varepsilon_{k_1}^- + \alpha\sigma)] + (d\mathbf{K}/dt)_0, \end{aligned} \quad (5)$$

where

$$|W_{k_1 k_2}|^2 = |V_{k_1 k_2}|^2 = (1/N^2) J^2 (|\mathbf{k}_1 - \mathbf{k}_2|) \quad (6)$$

and w_{sz} is the probability that S_z has the value S_z , and assuming the thermal equilibrium of the spin system,

$$w_{sz} = \exp\left(\frac{\alpha\sigma}{\kappa T} S_z\right) / \sum_{S_z = -S}^S \exp\left(\frac{\alpha\sigma}{\kappa T} S_z\right). \quad (7)$$

Further, at temperature lower than T_c , we have

$$\sigma = \sum_{S_z = -S}^S S_z w_{sz} \quad (8)$$

and $(d\mathbf{K}/dt)_0$ means the change due to other processes, for example, interactions with the phonons and impurities. As we are now interested in the anomalous resistance, and as the resistances due to each process are additive in our approximation method, we neglect the term $(d\mathbf{K}/dt)_0$ in the following calculations.

Next, we assume the distribution function n_k^\pm as Kramer's distribution

$$n_k^\pm = [\exp((\varepsilon_k^\pm - \varepsilon_0)/\kappa T - \mathbf{c} \cdot \mathbf{k}) + 1]^{-1}. \quad (9)$$

This assumption may be satisfactory when the interaction with the phonon is strong enough.

By the equation of the balancing of the total wave vector

$$(d\mathbf{K}/dt)_{field} + (d\mathbf{K}/dt)_{scatt} = 0, \quad (10)$$

we can determine \mathbf{c} .

§ 3. The calculations

To solve the equation of § 2, we use the following assumptions; the energy spectrum of the conduction electron can be written by using the effective mass m^* ; the Fermi energy is

much larger than the exchange energy $J(0)$ and the temperature: the crystal is isotropic, and, as usual, the summation \sum_k may be replaced by an integration over k .

The technique of the calculation is the same as that of ordinary electrical resistance. The calculation is straightforward and (10) becomes

$$n \frac{e}{\hbar} \mathbf{F} = c \frac{2\pi}{\hbar} \frac{3}{2} \frac{n^2}{N} \frac{m}{\hbar^2} \frac{\kappa T}{\mathcal{E}_0} (S - \sigma) (S + \sigma + 1) J_{eff}^2, \quad (11)$$

where k_0 is the wave vector of the Fermi surface, and

$$J_{eff}^2 = 4 \int_0^1 J^2(2k_0 x) x^3 dx. \quad (12)$$

On the other hand, the electric current \mathbf{j} is

$$\mathbf{j} = \mathbf{F} / \rho = (e/m) n \kappa T C \quad (13)$$

and thus the electrical resistance ρ is

$$\rho = (3\pi m^{*2} / N e^2 \hbar^2) (S - \sigma) (S + \sigma + 1) J_{eff}^2 / \mathcal{E}_0 \quad (14)$$

or introducing the ratio of the effective mass and the mass of electron,

$$\mu = m^* / m \quad (15)$$

and putting values of the physical constants,

$$\rho = 4.3 \times 10^{-4} \left(\frac{10^{23}}{N} \right) \mu^2 (S - \sigma) (S + \sigma + 1) \frac{J_{eff}^2}{\mathcal{E}_0} \text{ ohm} \quad (16)$$

where J_{eff} and \mathcal{E}_0 are measured in the unit of e.V. and N per c.c..

§ 4. Discussions

Equation (16) corresponds fairly well to the temperature dependence, as we expected. At absolute temperature zero, $\sigma = S$ and thus the electrical resistance $\rho = 0$. When the temperature is raised, σ decreases, and ρ increases. Above the Curie temperature, $\sigma = 0$, and ρ remains constant. To examine the temperature dependence more in detail, we should proceed to a higher order of the approximation. At sufficiently low temperatures, the method of the molecular field is not appropriate and we should use the spin-wave method. The situation is not simple unless we use the hypothesis that the spin-wave system is always in thermal equilibrium corresponding to Bloch's hypothesis for electron-phonon interaction. In the temperature near Curie temperature too, we should consider the short range ordering. For example, the fact that the curving of the resistance-temperature curve begins at a temperature slightly above the magnetic Curie point may be the effect of the short range ordering. Above the Curie temperature, the electrical resistance (16) is proportional to $S(S+1)$, thus in the metal of large spin, ρ would be very large. This tendency is seen also in the experimental results. For example, the anomalous electrical resistance of Fe is larger than that of Ni, and in Gd ($S=7/2$) particularly the anomalous electrical resistance is very large, and the ordinary resistance due to the electron-phonon interaction is rather small at room temperature. In rare earth metals, the anoma-

lous electrical resistance decreases when the spin value decreases, but in this case, except for Gd, there remains an orbital moment and our basic Hamiltonian (1) is not applicable. The calculation for such cases is now in progress.

Finally, we calculate the numerical values of the resistance at a temperature higher than T_0 for two cases, that is, Ni and Gd. In Ni, by Zener⁵⁾

$$N=0.9 \times 10^{23} \text{ per c.c., } S=0.3, \\ J_{eff} = J(0) = 0.48 \text{ e.v., } \varepsilon_0 = 3.14/\mu \text{ e.v.,} \quad (17)$$

and

$$\rho = 14 \times 10^{-6} \mu^3 \text{ ohm.} \quad (18)$$

The observed value is about 10×10^{-6} ohm, as shown in Fig. 1,
In Gd,

$$N=3.07 \times 10^{22} \text{ per c.c., } S=7/2, \\ J_{eff} = J(0) = 0.157 \text{ e.v., } \varepsilon_0 = 4.4/\mu \text{ e.v.,} \quad (19)$$

and

$$\rho = 120 \times 10^{-6} \mu^3 \text{ ohm} \quad (20)$$

where, to estimate $J(0)$, we have used the mean value of the results of the optical spectrum⁶⁾, multiplying by the same factor as in Ni. The experimental value is about 130×10^{-6} ohm.

In both cases, the theoretical and the experimental values coincide fairly well in the order of magnitude.

The author expresses his cordial thanks to Prof. Ariyama and also to Prof. Matsubara for their many valuable discussions.

References

- 1) For example, A. H. Wilson, *The Theory of Metals* pp. 273.
- 2) S. Legvold, F. H. Spedding, F. Earson and J. E. Elliot, *Rev. Mod. Phys.* **25** (1953), 229.
- 3) N. F. Mott, *Proc. Phys. Soc.* **47** (1935), 571; *Proc. Roy. Soc.* **153** (1936), 699; A. H. Wilson, *Proc. Roy. Soc.* **167** (1938), 580.
- 4) T. Kasuya, *Prog. Theor. Phys.* **16** (1956), 1.
- 5) C. Zener, *Phys. Rev.* **83** (1951), 259.
- 6) H. N. Russel, *J. Opt. Soc. Amm.* **40** (1950), 550.

Letters to the Editor

The Decay Interaction of $K_{\mu 3}$ and $K_{e 3}$ and Fermi Interaction

Susumu Furuichi, Tatsuro Kodama,
Yoshiko Sugahara and Minoru Yonezawa

*Department of Physics, Hiroshima University,
Hiroshima*

and

Akira Wakasa

*Department of Physics, Kanazawa University,
Kanazawa*

April 26, 1956

In the decay events $K_{\mu 3}$ and $K_{e 3}$ ¹⁾ for which we assume decay modes $K_{\mu 3(e 3)} \rightarrow \mu(e) + \nu + \pi^{0(\pm)}$, it is very interesting from the view point of the universality of Fermi interaction²⁾ to study what types of Fermi interaction may act in the decay interaction of $K_{\mu 3}$ and K_e , if the decays occur through the weak Fermi interaction (perhaps (A, N, μ, ν) or (Σ, N, μ, ν) in the $N-G-P$

scheme³⁾ via baryon loops. (see Fig. 1)

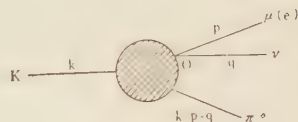


Fig. 1. The diagram for the decay of $K_{\mu^0(e^+)}$. Fermi interaction is acting at 0.

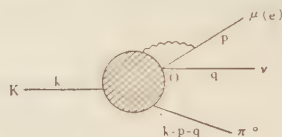


Fig. 2. One of the higher order corrections which are omitted in the present consideration. The waved line denotes the photon.

So in this note we show that it is possible to obtain some information about the types of Fermi interaction participating in these processes and also the type of K -meson from the energy spectra of the secondary muon and electron.

Following the method similar to that used by Dalitz⁵⁾ in the analysis of the $\tau \rightarrow$

Table I. Possible forms of $K_j F_j$

| Type of Fermi int. (O_j) Type of K -meson | $S(I)$ | $V(r_\mu)$ | $T(\sigma_{\mu\nu})$ | $A(r_5 r_\mu)$ | $P(r_5)$ |
|--|-----------------------|---|--|--|----------------------|
| S | forbidden | forbidden | $k_\mu(p+q)_\nu F^{S(T)}$ | $k_\mu F^{S(A)}$ $(p+q)_\mu F^{S(A)'} $ | $F^{S(F)}$ |
| V | forbidden | $\varepsilon_{\mu\nu\lambda\rho} k_\mu(p+q)_\nu F^{V(V)}$ | $k_\mu F^{V(T)}$ $(p+q)_\mu F^{V(T)'} $ | $\delta_{\mu\nu} F^{V(A)}$ | $(p+q)_\mu F^{V(P)}$ |
| PV | $(p+q)_\mu F^{PV(S)}$ | $\delta_{\mu\nu} F^{PV(V)}$ | $k_\mu F^{PV(T)}$ $(p+q)_\mu F^{PV(T)'} $ | $\varepsilon_{\mu\nu\lambda\rho} k_\mu(p+q)_\nu F^{PV(A)}$ | forbidden |
| PS | $F^{PS(S)}$ | $k_\mu F^{PS(V)}$ $(p+q)_\mu F^{PS(V)'} $ | $k_\mu(p+q)_\nu F^{PS(T)}$ | forbidden | forbidden |

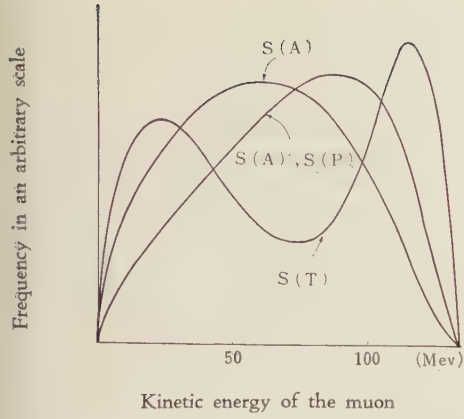


Fig. 3. Energy spectra of the secondary muon for the scalar K -meson. The histogram indicates the experimental data (14 examples).*

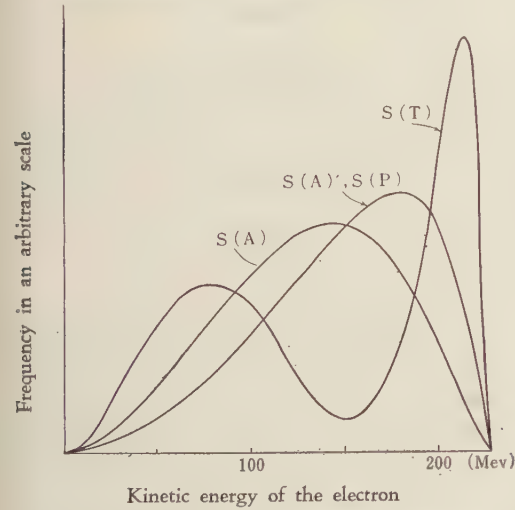


Fig. 5. Energy spectra of the secondary electron for the scalar K -meson. The histogram indicates the experimental data (11 examples).*

3π decay with the requirement of Lorentz invariance, we obtain the matrix elements of the general form

$$M = \bar{\psi}^\mu(p) O_j \psi^\nu(q) \phi^*(k-p-q) \times \Phi(k) K_j F^j(k, p, q).$$

Here $\Phi(k)$, $\phi(k-p-q)$, $\psi^\mu(p)$ and $\psi^\nu(q)$

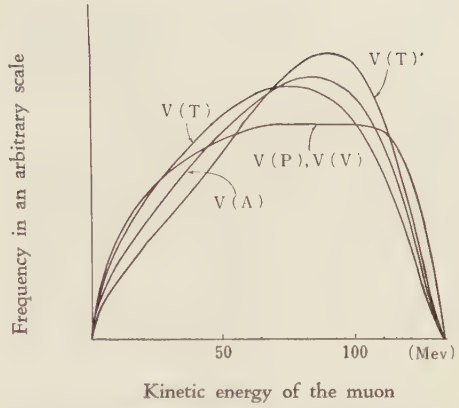


Fig. 4. Energy spectra of the secondary muon for the vector K -meson.

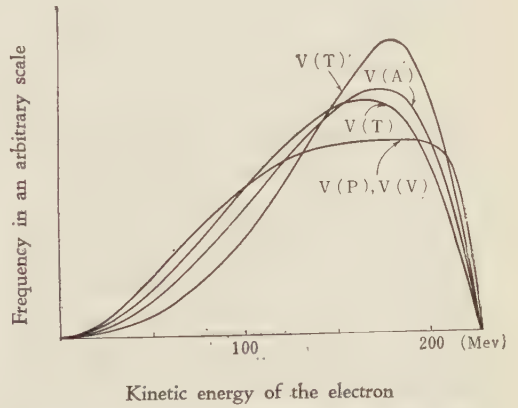


Fig. 6. Energy spectra of the secondary electron for the vector K -meson.

are the wave functions of the neutral pion, the muon (electron) and the neutrino respectively, and F^j 's may be scalar functions of 4-momenta k , p and q which are taken to be constants in our approximation. O_j 's are Dirac matrices.

Possible forms of $K_j F^j$'s lowest depen-

* Reported at Pisa Conference, and reference 2.

dent on four-momenta are tabulated in Table I, corresponding to *each* type of Fermi interaction. Here the spin of the K -meson is assumed to be 0 or 1.

Notice that since we assume Fermi interaction the muon momentum p and the neutrino momentum q appear only in the combination $p+q$, neglecting the small contribution from the higher order corrections such as is shown in the diagram in Fig. 2.

The energy spectra of the secondary muon or electron in the K -meson decays calculated from these matrix elements with the mass of K 965 m_e are shown in Figs. 3-6. The pseudoscalar K -meson and the pseudovector K -meson give the same spectra as the scalar K and the vector K respectively, only if the types of Fermi interaction are altered to the corresponding allowed types. And a similar situation occurs if we take the opposite parity for the muon (electron) to that for the neutrino.

At present the poor statistics allows us no definite conclusion, but the rapid increase of the number of observations is expected. Besides the energy spectra of the muon and the electron, the analysis along the same line of the angular distribution of $\hat{\mu}\pi^0$ and $\hat{e}\pi^0$ also gives more information of these decay interactions.

The detailed discussion of this analysis will shortly be published in a later issue of this journal.

The authors are indebted to Dr. S. Ogawa for his suggestion of the theme and helpful discussions. They also thank Prof. K. Sakuma, Prof. T. Tati, Dr. S. Oneda and Dr. S. Hori for their kind interest in this work.

- 1) M. G. Menon and C. O'Ceallaigh, *Proc. Roy. Soc. (London)* **221** (1954), 292.
- 2) M. F. Kaplon, J. Klarmann and G. Yekutieli, *Phys. Rev.* **99** (1955), 1528.
- 3) E. J. Konopinski and H. M. Mahmoud, *Phys. Rev.* **92** (1953), 1045. See also K. Iwata, S. Ogawa, H. Okonogi, B. Sakita and S. Oneda, *Prog. Theor. Phys.* **13** (1955), 19.
- 4) K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285; M. Gell-Mann and A. Pais, *Proceedings of the Glasgow Conference 1954* (Pergamon Press, London, 1955).
- 5) R. H. Dalitz, *Phil. Mag.* **44** (1953), 1068.

Interaction of K^+ -Meson and Nucleon

Shigeo Minami

*Department of Physics, Osaka City University,
Osaka*

May 16, 1956

Recently Lannutti et al.¹⁾ have observed the interactions of K^+ -mesons in flight in nuclear emulsions. According to their estimation there were four charge exchange and 36 non-charge exchange events in the energy region of 30~120 Mev.

In this paper we intend to show some relation between the ratio of charge exchange to non-charge exchange cross sections and the relative parities of hyperons and heavy mesons. For simplicity, we assume that the spins both of Λ and of Σ are 1/2 and that the spin of K^+ is 0, then we can write the interaction Hamiltonian as follows:^{2)*}

* Of course this interaction satisfies the well-known Gell-Mann³⁾ Nishijima's⁴⁾ rule.

$$H_{N\Lambda K} = f\bar{\psi}_\Lambda O\phi_K^* \tau_4 \psi + \text{h.c.}, \quad (\tau_4 = 1),$$

$$H_{N\Sigma K} = g\bar{\Psi}_\alpha O'\phi_K^* \tau_\alpha \psi + \text{h.c.}, \quad (1)^*$$

where ψ_Λ , Ψ , ψ and ϕ_K are Λ , Σ , nucleon and K -meson fields respectively, and

$$\left\{ \begin{array}{l} O \text{ (or } O') = 1 \text{ for spinor hyperon and} \\ \text{scalar heavy meson, or} \\ \text{for pseudospinor hyperon} \\ \text{and pseudoscalar heavy} \\ \text{meson,} \\ O \text{ (or } O') = i\gamma_5 \text{ for spinor hyperon and} \\ \text{pseudoscalar heavy meson,} \\ \text{or for pseudospinor hy-} \\ \text{peron and scalar heavy} \\ \text{meson.} \end{array} \right.$$

Now, let us calculate the ratio of charge exchange to non-charge exchange scattering

$$1/R = \sigma(K^+ + n \rightarrow K^0 + p) / \{ \sigma(K^+ + p \rightarrow K^+ + p) + \sigma(K^+ + n \rightarrow K^+ + n) \} \quad (2)$$

by means of perturbation theory. As the angular distribution for these reactions at 50 Mev (lab.) can be expressed approximately by the form of $(a + b \cos \theta)$ in the center-of-mass system, we estimate this ratio by the differential cross sections at 90° and obtain the following results.

$$1/R = (C_1 - 2C_2x + C_3x^2) / (C_1 + 2C_2x + 5C_3x^2), \quad (3)$$

where $x = g^2/f^2$ and C 's are given in Table 1.

Table 1

| Case | O | O' | C ₁ | C ₂ | C ₃ |
|-------|-------------|-------------|----------------|----------------|----------------|
| (I) | 1 | 1 | 6.91 | 6.26 | 5.68 |
| (II) | 1 | $i\gamma_5$ | 6.91 | -3.04 | 1.34 |
| (III) | $i\gamma_5$ | 1 | 1.46 | -2.88 | 5.68 |
| (IV) | $i\gamma_5$ | $i\gamma_5$ | 1.46 | 1.40 | 1.34 |

From eq. (3),

$$C_3(R-5)x^2 - 2C_2(R+1)x + C_1(R-1) = 0. \quad (4)$$

Here, it should be noticed that there are the following allowable ranges for R because $x = g^2/f^2$ must be a positive number.

$$\left\{ \begin{array}{ll} R \geq 1 & \text{for the case of (I) or (IV),} \\ 5 > R \geq 1/2 & \text{for the case of (II) or (III).} \end{array} \right. \quad (5)$$

In spite of the energy dependence of C 's, these allowable ranges are not affected by the energy, at least, up to 100 Mev.

Although the value $R=9$ which has been given by Lannutti et al. may suffer some changes by the increasing accuracy of experimental data, it may be said that this value is consistent with the result derived from (I) or (IV) of our model.*

Since we adopt a special model and employ the perturbation theory, it will be difficult to reach any definite conclusion. But, as mentioned above, the value of R is sensitive to the relative parities of hyperons and heavy mesons, so we may expect that R will play an important role for determining the parities of strange particles.

In conclusion the author should like to express his thanks to Prof. Y. Yamaguchi and Mr. S. Iwao for their valuable discussions.

- 1) J. E. Lannutti, W. W. Chupp, G. Goldhaber, S. Goldhaber, E. Helmy, E. L. Iloff, A. Pevsner and D. M. Ritson, Phys. Rev. **101** (1956), 1617.

* More detailed accounts of the analysis of $K^+ \rightarrow$ nucleon scattering will be published in near future.

- 2) C. Iso, Soryushiron Kenkyu (Mimeographed Circular in Japanese) **11** (1956), 1.
S. Iwao, Soryushiron Kenkyu (Mimeographed Circular in Japanese) **11** (1956), 240.
- 3) M. Gell-Mann and A. Pais, Proceeding of the 1954 Glasgow Conference.
- 4) T. Nakano and K. Nishijima, Prog. Theor. Phys. **10** (1953), 581.
K. Nishijima, Prog. Theor. Phys. **13** (1955), 285.

Nature of Nuclear Forces Indicated by the Photodisintegration of the Deuteron, II

Shih-Hui Hsieh

Physical Institute, Nagoya University, Nagoya

May 24, 1956

In the previous letter¹⁾ we have proposed a method to observe properties of nuclear forces. It is to analyze the photodisintegration of the deuteron using the phase shift analysis of p - p scattering. Using this method some interesting features of nuclear forces can be brought to light. One of these will be shown here.

In the photodisintegration of the deuteron it seems to be well established that the *isotropic part is large* in a wide range of energies, although the experiments are not very accurate yet.²⁾ We wish to point out that this fact can be accounted for, if δ_0^0 is smaller than, and if δ_1^0 , δ_1^1 , δ_1^2 are larger than and are quite different from what they have been thought up to today. Here we only show the result of calculation at 11.2 Mev referring to the phase shift analysis of 18 Mev p - p scattering.

The p - p scattering at 18 Mev was investigated by Yntema and White,³⁾ and

the phase shift analysis was done by Clementel et al.⁴⁾ Because of the experimental error the numerical values given by Clementel et al. is not quite definite. However it is possible to make use of these data to reveal the properties of nuclear forces mentioned above.

They interpreted p - p scattering data by means of four phase shifts δ_0^0 , δ_1^0 , δ_1^1 , and δ_1^2 . For a given angular distribution we can determine δ_0^0 at first, and then corresponding to one value of δ_0^0 three sets of the phase shifts, δ_1^1 , δ_1^1 , and δ_1^2 are obtained.

In our case for 18 Mev p - p scattering, the following result was given by Clementel et al. 1) If one assumes that $\sigma(30^\circ) = 25$ mb, $\sigma(40^\circ) = 26.60$ mb, and $\sigma(60^\circ) = 27.37$ mb, one finds $\delta_0^0 = 43.7^\circ$. 2) Assuming that the contribution from the nuclear part is isotropic, and $\sigma(90^\circ) = 27.37$ mb, δ_0^0 is determined to be 50.5° . For each δ_0^0 three sets of δ_1^0 , δ_1^1 , δ_1^2 are given as shown in Table 1.

Table 1. Phase shifts which can account for 18 Mev p - p scattering given by Clementel et al. The first three sets are obtained by assuming $\sigma(30^\circ) = 25$ mb, $\sigma(40^\circ) = 26.60$ mb, and $\sigma(60^\circ) = 27.37$ mb. Other three sets are for the case when the nuclear part is assumed to be isotropic. The last column indicates the sign of polarization.

| | δ_0^0 | δ_1^0 | δ_1^1 | δ_1^2 | P |
|---|--------------|--------------|--------------|--------------|-----|
| A | 43.7° | 19.4° | -7.4° | 2.75° | + |
| B | 43.7° | -14.6° | 10.5° | -1.5° | - |
| C | 43.7 | 24.1° | -4.3° | -1.1° | - |
| D | 50.5° | 3.3° | -4.3° | 1.5° | + |
| E | 50.5° | -4.8° | 3.6° | 1.9° | + |
| F | 50.5° | 6.3° | -3.2° | 0.25° | + |

Now for the calculation of the photodisintegration of the deuteron⁵⁾ referring to these data, the following procedure was

adopted.

As the wave function of the deuteron we used that of Feshbach and Schwinger,⁽⁶⁾ whose numerical formula was given by Austern.⁽⁷⁾ The effect due to the detailed form of the wave function of the deuteron is very small at low energies.

Since σ_{mq} (σ_{mq} indicates the magnetic quadrupole cross section and so on.) and σ_{eq} are quite negligible,⁽⁸⁾ we calculated only σ_{md} and σ_{ed} . Also we neglected the contribution from the exchange current in the calculation of σ_{md} , for its contribution is very small at such low energies.⁽⁹⁾

In the calculation of σ_{ed} (the final states are 3P_0 , 3P_1 , and 3P_2 .) we used free 3P waves shifted by δ_1 in Table 1 at all inter-nuclear distances (because the contribution from $r < 1.0/\kappa^{(10)}$ to the required integral is very small).

For σ_{md} we estimated its upper and lower limits. The upper limit is given, when we take 1S_0 wave in the form

$$U_0 = \sin(kr + \delta_0^0), \quad (1)$$

and the lower limit is given, when we take 1S_0 wave in the form

$$U_0 = \sin(kr + \delta_0^0) (1 - e^{-\zeta r}), \quad (2)$$

where ζ is determined from the effective range theory ($\zeta = 1.59 \kappa$). In the first case σ_{md} is estimated somewhat larger, and in the second case somewhat smaller. It is to be mentioned here that the effect of a hard core is very small at low energies.

The result is shown in Table 2. The angular distribution of the photodisintegration of the deuteron can be expressed in the form

$$a + b \sin^2 \vartheta, \quad (3)$$

if we neglect the contribution from σ_{mq}

and σ_{eq} . According to Allen's estimation,⁽¹¹⁾

$$a/b = 0.05 \pm 0.02 \quad (4)$$

at this energy. From Table 2 we see that, for the three cases the calculated value of a/b falls within the experimental error given in (4), when $\delta_0^0 = 43.7^\circ$. However when we

Table 2. The calculated result of the photodisintegration of the deuteron at 11.2 Mev using the phase shifts given in Table 1. The upper half of the table is for the case when we estimate σ_{md} by its upper limit, and the lower half is for the case when σ_{md} is estimated by its lower limit. All cross sections are in unit of 10^{-29} cm^2 .

| | σ_{ed} | σ_{md} | σ_T | σ_{ed}^{ip} | σ_{md}^{ip} | σ_{ip} | a/b |
|---|---------------|---------------|------------|--------------------|--------------------|---------------|-------|
| A | 108 | | 113 | 2.08 | | 6.38 | 0.040 |
| B | 115 | 4.84 | 120 | 4.79 | 4.30 | 9.09 | 0.055 |
| C | 107 | | 112 | 2.48 | | 6.78 | 0.043 |
| D | 108 | | 112 | 1.04 | | 4.60 | 0.028 |
| E | 115 | 4.05 | 119 | 1.76 | 3.56 | 5.32 | 0.030 |
| F | 109 | | 113 | 0.93 | | 4.49 | 0.028 |
| A | 108 | | 111 | 2.08 | | 4.53 | 0.029 |
| B | 115 | 2.72 | 118 | 4.79 | 2.45 | 7.24 | 0.044 |
| C | 107 | | 110 | 2.48 | | 4.93 | 0.031 |
| D | 108 | | 110 | 1.04 | | 2.90 | 0.018 |
| E | 115 | 2.11 | 117 | 1.76 | 1.86 | 3.62 | 0.021 |
| F | 109 | | 111 | 0.93 | | 2.79 | 0.017 |

adopt $\delta_0^0 = 50.5^\circ$, a/b turns out to be too small. Therefore the smaller δ_0^0 is favourable.

Now we wish to discuss the reason revealing the nature of nuclear forces. In Table 2 we see that, if δ_0^0 is smaller, σ_{md}^{ip} (the isotropic part of σ_{md}) is larger. This can be understood as follows. The main part of σ_{md}^{ip} is due to $\int u_0 u_d dr^{(12)}$ (where U_d is the 3S_1 part of the wave function of the deuteron), and this integral is larger

when δ_0^0 is smaller. The reason is that, if δ_0^0 is small, U_0 is pushed out toward large r , and therefore makes the integrated value large.

Next if δ_0^0 is smaller, the three phase shifts δ_1^0 , δ_1^1 , and δ_1^2 must be larger in their absolute value, and, moreover, the difference between any two of them must be larger, in order to account for p - p scattering. Such a tendency is just what we require, because the large and separated values of δ_1^0 , δ_1^1 , and δ_1^2 make σ_{el}^{ip} (the isotropic part of σ_{el}) larger. (See Table 2). This point can be seen from the formula of σ_{el}^{ip} , which is given by Rarita and Schwinger.¹²⁾

Therefore in such cases, where δ_0^0 is small and δ_1^0 , δ_1^1 , δ_1^2 are large and different from each other, both σ_{mut}^{ip} and σ_{el}^{ip} are increased, and make σ_{ip} large. This is favourable to the photodisintegration of the deuteron.

Our conclusion can also be summarized in the following way. *To account for both photodisintegration of the deuteron and the p - p scattering, we must adopt a smaller value of δ_0^0 , and larger values of δ_1^0 , δ_1^1 , and δ_1^2 , than they have usually been treated up to today.* Similar conclusion has been obtained by J. Iwadare et al.¹³⁾ by a different analysis (from the point of view of the pion theory).

At last we wish to point out that, improved experiments on photodisintegration of the deuteron and the polarization of p - p scattering will determine more reliable values of the four phase shifts, δ_0^0 , δ_1^0 ,

δ_1^1 and δ_1^2 . If the polarization is positive at 18 Mev, the first set of Table 1 is the best among the six sets.

The author wishes to express his thanks to Mr. S. Otsuki for his valuable discussions.

- 1) S. Hsieh and M. Nakagawa, Prog. Theor. Phys. **15** (1956), 79. All notations used in this letter are the same as those in this reference.
- 2) E. A. Whalin, Phys. Rev. **95** (1954), 1362; Lew Allen, Phys. Rev. **98** (1955), 705.
- 3) J. L. Yntema and M. G. White, Phys. Rev. **95** (1954), 1226.
- 4) Beretta, Clementel, and Villi, Nuor. Cim. **1** (1955), 739.
- 5) For the formulas on the photodisintegration of the deuteron see W. Rarita and J. Schwinger, Phys. Rev. **59** (1941), 436; N. Austern, Phys. Rev. **85** (1952), 283.
- 6) H. Feshbach and J. Schwinger, Phys. Rev. **84** (1951), 194.
- 7) N. Austern, Phys. Rev. **85** (1951), 283.
- 8) J. F. Marshall and E. Guth, Phys. Rev. **84** (1950), 738.
- 9) Any theory of the exchange current which has been published up to today predicts a small contribution to σ_{mut} due to the exchange current at such low energies. See F. Villars, Phys. Rev. **72** (1947), 256; N. Austern and G. Sachs, Phys. Rev. **81** (1951), 710; N. Austern, Phys. Rev. **85** (1952), 283; A. Sugie and S. Yoshida, Prog. Theor. Phys. **10** (1953), 236; M. Sugawara, Prog. Theor. Phys. **14** (1955), 535; S. Hatano, Prog. Theor. Phys. **14** (1955), 170 etc.
- 10) $\kappa = m_\pi c/\hbar$, and $1/\kappa = 1.40 \times 10^{-13}$ cm.
- 11) Lew Allen, Private communication. The author thanks to Dr. Allen for sending him experimental data.
- 12) See 5).
- 13) J. Iwadare, S. Otsuki, M. Sano, S. Takagi and W. Watari. To be published.

On the Quantum Mechanics-like Description of the Theories of the Brownian Motion and Quantum Statistical Mechanics*

Nobuhiko SAITÔ and Mikio NAMIKI

Department of Applied Physics, Waseda University, Tokyo

(Received March 30, 1956)

Analogies among the theories of Brownian motion, quantum mechanics and quantum statistics are pointed out from various aspects. Along this line of thought quantum mechanics-like descriptions of Brownian motion and quantum statistics are presented.

§ 1. Introduction

In quantum mechanics, of essential importance is the introduction by Born of the notion of probability for the square of the absolute value of the wave function. The role of probability in quantum mechanics lies in this respect and the wave function itself does not have any direct physical meaning. On the other hand, in connection with wave function and probability theory, Schrödinger¹⁾ and Metadier²⁾ pointed out that the Schrödinger equation for a free particle

$$-(\hbar/i)\partial\psi/\partial t = -\hbar^2/2m \cdot \Delta\psi \quad (1.1)$$

and the diffusion equation which appears in the Einstein's theory of the Brownian motion

$$\partial W/\partial t = D\Delta W \quad (1.2)$$

have a similar form and $(\hbar/2m)i$ is to be compared with diffusion constant D . Furthermore Fürth³⁾ showed that the measured values of positions and velocities of a Brownian particle give rise to an uncertainty relation

$$\langle x^2 \rangle \langle v^2 \rangle \gtrsim D^2, \quad (1.3)$$

which again gives the same correspondence between \hbar/m and D as above. These facts suggest that there will exist a close formal analogy in the structures of both theories, although, as is said above, in the theory of Brownian motion the function W has the meaning of probability, while in quantum mechanics ψ itself is not the probability, but only the probability amplitude. Thus the connection of quantum mechanics to probability theories is twofold, i.e., firstly for the square of the absolute value of wave function and secondly for the evolution of wave function as a stochastic process. In this note we start

* A preliminary report of this paper was published in *Busseiron Kenkyû*, No. 78 (1954) 95 (in Japanese).

to elucidate in detail the latter properties of quantum mechanics and quantum statistics, comparing with those of the Brownian motion.

§ 2. Analogies in the theories of Brownian motion, quantum mechanics and quantum statistics

Brownian motions are regarded as stationary Markoff processes. Let $T(x_2, t_2|x_1, t_1)$ be the conditional probability density or transition probability that a Brownian particle which was at x_1 at time t_1 lies at x_2 at t_2 , then we have the following Kolmogoroff-Chapman's equation

$$T(x_2, t_2|x_1, t_1) = \int T(x_2, t_2|x_3, t_3) T(x_3, t_3|x_1, t_1) dx_3$$

$$t_2 > t_3 > t_1. \quad (2.1)$$

This was the starting point of Einstein's theory of the Brownian motion and a diffusion, or, in general* a Fokker-Planck equation is derived under some assumptions from eq.

* In the most general treatment, we have two kinds of Fokker-Planck equations. For the sake of simplicity, we confine ourselves here to one-dimensional problem.

Putting

$$\begin{aligned} T(x_2, t_2|x_1, t_1) &\equiv u(x_2, t_2) \\ &\equiv v(x_1, t_1) \end{aligned} \quad (2.1')$$

or
we have

$$\begin{aligned} \partial u(x_2, t_2)/\partial t_2 &= -\partial \{a(x_2) u(x_2, t_2)\}/\partial x_2 \\ &\quad + \partial^2 \{b(x_2) u(x_2, t_2)\}/\partial x_2^2 \\ &\equiv L_1 u(x_2, t_2) \end{aligned} \quad (2.2')$$

and

$$\begin{aligned} \partial v(x_1, t_1)/\partial t_1 &= -a(x_1) (\partial v(x_1, t_1)/\partial x_1) - b(x_1) (\partial^2 v(x_1, t_1)/\partial x_1^2) \\ &\equiv -L_2 v(x_1, t_1) \end{aligned} \quad (2.3')$$

where $a(x)$ and $b(x)$ are functions related respectively to the first and second moments of the displacement. These two functions u and v are so introduced as to describe the time variations of final state and initial state respectively. The differential operator L_1 is adjoint to L_2 , i.e., one has for two duely defined arbitrary functions f and g ,

$$\int g^* L_1 f dx = \int (L_2 g)^* f dx. \quad (2.4')$$

Let the eigenfunctions and eigenvalues of the operator L_1 be $u_m(x)$ and λ_m , then we have

$$L_1 u_m(x) = \lambda_m u_m(x) \quad (2.5')$$

$$L_2 v_m(x) = \lambda_m^* v_m(x) \quad (2.6')$$

and

$$\int u_m^*(x) v_n(x) dx = \delta_{nm} \quad (2.7')$$

where v_m is the eigenfunction of the operator L_2 belonging to the eigenvalue λ_m^* . The eigenvalue λ_m is real ($\lambda_m = \lambda_m^*$) and its inverse $1/\lambda_m$ is called relaxation time. In effect the transition probability is easily shown to be developed as follows,

$$T(x_2, t_2|x_1, t_1) = \sum_n u_n(x_2) e^{-\lambda_n t} v_n^*(x_1), \quad t = t_2 - t_1. \quad (2.8')$$

In quantum mechanics the differential operator is self-adjoint or hermitian. The discussions thereof are simpler than those of Brownian motion. The probability density $W(x_2, t_2)$ that the Brownian particle lies at x_2 at t_2 satisfies the differential equation (2.2'), and this is the so-called Fokker-Planck equation.

(2.1), e.g., in Einstein's case of a free particle one obtains the equation for free diffusion (1.2).

In quantum mechanics let $\psi(x, t)$ be the wave function and we have

$$\psi(x_2, t_2) = \int K(x_2, t_2 | x_1, t_1) \psi(x_1, t_1) dx_1, \quad (2.2)$$

where $K(x_2, t_2 | x_1, t_1)$ is the transformation function. The latter function itself obeys a Kolmogoroff-Chapman's equation,

$$K(x_2, t_2 | x_1, t_1) = \int K(x_2, t_2 | x_3, t_3) K(x_3, t_3 | x_1, t_1) dx_3. \quad (2.3)$$

This implies nothing but that quantum mechanics can be described formally as a Markoff process and its Fokker-Planck equation is the Schrödinger equation.*

From eq. (1.2) Einstein obtained for the transition probability the following expression,

$$T(x_2, t_2 | x_1, t_1) = \frac{1}{\{4\pi D(t_2 - t_1)\}^{3/2}} \exp\left\{-\frac{(x_2 - x_1)^2}{4D(t_2 - t_1)}\right\}, \quad (2.4)$$

and later Wiener⁴⁾ gave a mathematical formulation to Einstein's theory. He rewrote Kolmogoroff-Chapman's equation (2.1) using (2.4) and introduced an idea of path integral, giving a measure for the path; that is, he expressed the transition probability as follows,

$$T(x_2, t_2 | x_1, t_1) = N \int \exp\left\{-\frac{1}{4D} \int_{t_1}^{t_2} \left(\frac{dx}{dt}\right)^2 dt\right\} d(\text{path}) \quad (2.5)$$

where $\int_{t_1}^{t_2} (dx/dt)^2 dt$ is an integral along a path, and $\exp\{\dots\}$ corresponds to the above-mentioned measure, with which the integral $\int \dots d(\text{path})$ is carried out. N is the normalization constant. Path integrals as above are called Wiener integrals. In Einstein's theory, for the projection q of x upon an axis we have from eq. (2.4)

$$\langle (q_2 - q_1)^2 \rangle = 2D(t_2 - t_1), \quad (2.6)**$$

* In the theory of Brownian motion we have a normalization condition

$$\int W(x, t) dx = 1,$$

on the other hand in quantum mechanics normalization condition holds not for ψ itself but for $|\psi|^2$.

$$\int |\psi(x, t)|^2 dx = 1.$$

This yields a trivial difference in form between Schrödinger equation and general Fokker-Planck equation.

** From eq. (2.6) we have for the velocity

$$\langle \Delta v^2 \rangle = \langle \Delta q^2 \rangle / t^2 = 2D/t,$$

therefore the uncertainty relation (1.3) is rather written as

$$\langle \Delta v^2 \rangle \langle \Delta q^2 \rangle \simeq 4D^2$$

or

$$\langle \Delta p \rangle \langle \Delta q \rangle \simeq 2mD$$

where p is the momentum.

therefore the velocity of Brownian particle which corresponds to the time derivative of q can not be defined. The stochastic process like this is called Wiener process. Einstein's consideration lies in that the motion of a Brownian particle can never be described by the classical deterministic dynamics, but must be treated as a stochastic process. On the other hand Langevin settled an equation of motion of a Brownian particle, as follows,

$$m(d^2\mathbf{x}/dt^2) + \zeta(d\mathbf{x}/dt) = \mathbf{X}(t)$$

where ζ is the friction constant, $\mathbf{X}(t)$ is the fluctuating force which gives rise to the irregular motion of the particle. $\mathbf{X}(t)$ is assumed to have the properties* that its average (over an ensemble or over a long time) is zero and there is no correlation between $\mathbf{X}(t)$'s at different times. Then one can easily integrate the above equation (2.7) and show that Einstein's relation (2.6) holds only after a long time, and in short time interval t one has

$$\langle (\mathbf{x}_2 - \mathbf{x}_1)^2 \rangle \simeq v_0^2 t^2, \quad (2.8)$$

where v_0 is the initial velocity; this is because the influence of the velocity of the immediately preceding past persists in short period. In this case the velocity can be defined and this stochastic process is called an O—U (Ornstein-Uhlenbeck) process. The path integral formulation for the O—U process was given by Onsager-Machlup⁽⁵⁾ and Hashitsume⁽⁶⁾ and was applied to derive a variational principle in the irreversible processes.

In quantum mechanics the path integral formulation was introduced by Feynman⁽⁷⁾ in his space-time approach to quantum mechanics as follows,

$$K(\mathbf{x}_2, t_2 | \mathbf{x}_1, t_1) = N \int \exp \{ (i/\hbar) I \} d(\text{path}), \quad (2.9)$$

$$I = \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}, t) dt, \quad (2.10)$$

where L is the Lagrangian of the system and I is the action integral. This formulation is based on the Kolmogoroff-Chapman's relation (2.3). A similar consideration of Wiener integral formulation in quantum mechanics was discussed by Montroll.⁽⁸⁾

The density matrix which plays a fundamental role in quantum statistics is written as

$$\rho(\mathbf{x}_2 | \mathbf{x}_1; \beta) = \sum_n \phi_n(\mathbf{x}_2) e^{-\beta E_n} \phi_n^*(\mathbf{x}_1), \quad \beta = 1/kT, \quad (2.11)$$

where ϕ_n is the eigenfunction belonging to eigenvalue E_n . For the sake of convenience we define β_2' and β_1' by $\beta_2' - \beta_1' = 1/T$ and write the density matrix as $\rho(\mathbf{x}_2, \beta_2' | \mathbf{x}_1, \beta_1')$, then we can show easily

$$\rho(\mathbf{x}_2, \beta_2' | \mathbf{x}_1, \beta_1') = \int \rho(\mathbf{x}_2, \beta_2' | \mathbf{x}_3, \beta_3') \rho(\mathbf{x}_3, \beta_3' | \mathbf{x}_1, \beta_1') d\mathbf{x}_3. \quad (2.12)$$

In other words this is a Kolmogoroff-Chapman's equation, if β' is regarded as time. The Fokker-Planck equation for (2.12) is the well-known Bloch equation

* A more detailed qualification for \mathbf{X} is discussed in § 3.

$$-k(\partial\rho/\partial\beta')=H\rho. \quad (2.13)$$

Comparing this equation with Schrödinger equation

$$-(\hbar/i)\partial\psi/\partial t=H\psi, \quad (2.14)$$

a correspondence

$$it/\hbar \rightarrow \beta'/k (= \beta) \quad (2.15)$$

exists here. Rather one may say that Schrödinger equation for transformation function

$$-(\hbar/i)\partial K(x_2, t_2|x_1, t_1)/\partial t_2 = H(x_2, t_2) K(x_2, t_2|x_1, t_1) \quad (2.16)$$

corresponds to Bloch equation (2.13). Bloch⁽⁹⁾ applied this formalism to the theory of ferromagnetism. Recently Feynman⁽¹¹⁾ discussed the problem of liquid helium by introducing path integral formulation for density matrix using the correspondence (2.15). The transformation function is developed as follows,

$$K(x_2, t_2|x_1, t_1) = \sum \phi_n(x_2) e^{-iE_n(t_2-t_1)/\hbar} \phi_n^*(x_1). \quad (2.17)$$

This development shows that the transformation function and density matrix have same structure with correspondence (2.15).

The analogue in the theory of the Brownian motion of (2.11) in density matrix and (2.17) in quantum mechanics is the one encountered in the method of eigenfunction development frequently used in the theory of relaxation phenomena. The essentials are described in the foot-note to § 2 (p. 2).

Above considerations suggest that there is a close similarity in the formal structure of the theories of Brownian motion, quantum mechanics and quantum statistics. They are all described formally as stationary Markoff processes. Table I shows the analogies of these theories.

Table I.

| Brownian Motion | Quantum Mechanics | Quantum Statistics |
|--|--|---|
| Transition probability $T(x_2, t_2 x_1, t_1)$ | Transformation function $K(x_2, t_2 x_1, t_1)$ | Density matrix $\rho(x_2, \beta_2' x_1, \beta_1')$ |
| Kolmogoroff-Chapman's equation eq. (2.1) | eq. (2.3) | eq. (2.12) |
| Fokker-Planck equation | Schrödinger equation | Bloch equation |
| Eigenfunction expansion $T = \sum u_n e^{-\lambda_n t} v_n^*$ | $K = \sum \phi_n e^{-iE_n t/\hbar} \phi_n^*$ | $\rho = \sum \phi_n e^{-\beta E_n} \phi_n^*$ |
| Wiener integral | Feynman integral | Feynman integral |
| Uncertainty relation (Wiener process) $\langle \Delta p \rangle \langle \Delta q \rangle \simeq 2mD$ | $\langle \Delta p \rangle \langle \Delta q \rangle \simeq \hbar$ | |

On the other hand one may say that the theory of quantum mechanics is one of the most beautiful and magnificent theories achieved by human race, and the various techniques

have been invented for the wide problems of quantum mechanics. The above mentioned analogies permit us to apply the techniques of quantum mechanics to the theories of Brownian motion and quantum statistics. These are discussed in detail in the following sections.

§ 3. "Schrödinger" representation of the Brownian motion

Schwinger's dynamical principle of quantum mechanics has an advantage that the various relations which are obtained separately in the traditional theories of quantum mechanics are derived from one variational principle. For example we obtain Schrödinger equation, Heisenberg equation or commutation relation, etc., from this principle. In Appendix we give a review of Schwinger's dynamical principle in a manner somewhat different from the original description of Schwinger. This method can be applied immediately to the theories of Brownian motion and quantum statistics, provided that we can write down the path integral formulation for transition probability or density matrix. For this purpose we will give a brief discussion for the method of introduction of path integral in the theory of Brownian motion.

Langevin equation for a Brownian particle under external force with potential $U(x)$ is

$$m\ddot{x} + \zeta\dot{x} + U'(x) = X(t), \quad (3.1)$$

where $X(t)$ is the fluctuation force, which is assumed to have the following properties

$$\text{i) } \langle X(t) \rangle = 0 \quad (3.2)$$

$$\text{ii) } \langle X(t_1) X(t_2) \rangle = \langle X^2 \rangle \delta(t_1 - t_2) \quad (3.3)$$

and $\langle X^2 \rangle$ is independent of time.

iii) Further the probability density for X is given by

$$\Psi(X) \propto \exp(-X^2/4\zeta kT). \quad (3.4)$$

This Gaussian form is easily suggested by central limit theorem from the fact that the fluctuation force is regarded as a resultant of numerous elementary random forces due to impacts of molecules, and the coefficient $4\zeta kT$ is determined so that the distribution of the velocity of the Brownian particles should lead to Maxwellian after sufficiently long time.¹¹⁾

Define

$$B_t(\Delta t) = \int_t^{t+\Delta t} X(t) dt \simeq X(t) \Delta t, \quad \text{for small } \Delta t, \quad (3.5)$$

and we have from (3.3)

$$\langle B_t(\Delta t)^2 \rangle = \iint_t^{t+\Delta t} \langle X(t) X(s) \rangle dt ds = \langle X^2 \rangle \Delta t, \quad (3.6)$$

Then we can easily show that the distribution of $B_i(\Delta t)$ is Gaussian, i.e., for the probability density of $B_i(\Delta t)$

$$\Psi \{B_i(\Delta t)\} \propto \exp \{-B_i^2(\Delta t)/4\zeta kT\Delta t\}. \quad (3.7)$$

On the other hand for small time interval between t and $t + \Delta t$, the path of a Brownian particle is determined by specifying $X\Delta t = B_i(\Delta t)$, whose probability density is given by (3.7). Consequently the probability of a path from t_1 to t_2 is given from ii) by

$$\Psi \{B_{t_1}(\Delta t)\} \Psi \{B_{t_1+\Delta t}(\Delta t)\} \cdots \Psi \{B_{t_2-\Delta t}(\Delta t)\} \\ \propto \exp \left\{ - (1/4\zeta kT\Delta t) \sum B_i^2(\Delta t) \right\} \xrightarrow{\Delta t \rightarrow 0} \exp \left\{ - (1/4kT\zeta) \int_{t_1}^{t_2} X(t)^2 dt \right\}. \quad (3.8)$$

That is to say, if $X(t)$ for $t_1 \leq t \leq t_2$ and initial conditions are given, the path of a Brownian particle is determined completely. Eq. (3.8) gives the probability for a path. Consequently, the transition probability is written down at once, as follows

$$T(\text{final state}; \text{initial state}) \propto \int \exp \left\{ - (1/4\zeta kT) \int X^2 dt \right\} d(\text{path}), \quad (3.9)$$

which is the form of path integral given by Onsager and Machlup.⁵⁾ (See also Hashitsume^{11a)}

The above formulation is for O—U processes. For Wiener processes, we have only to neglect the inertia term in Langevin equation. Then Langevin equation (3.1) reduces to

$$\dot{x} + (1/\zeta) U'(x) = (1/\zeta) X(t) \equiv Y(t), \quad (3.10)$$

and the transition probability turns out to be

$$T(\text{final} | \text{initial}) \propto \int \exp \left\{ - (\zeta/4kT) \int_{t_1}^{t_2} Y(t)^2 dt \right\} d(\text{path}). \quad (3.11)$$

This is the path integral representation of Wiener for Brownian motion which reduces to (2.5) for free particles ($U(x)=0$).

The important difference between the two processes, Wiener and O—U, is that the initial condition to be required for the determination of a path is the values of position and velocity in an O—U process, and the value of position in a Wiener process. Consequently the initial and final states in the transition probability are specified by positions and velocities in an O—U process and by positions in a Wiener process. In the present section, a few examples will be discussed in a manner analogous to the “Schrödinger” representation in quantum mechanics. The “Heisenberg” representation is treated in § 4.

3.1 Wiener process

As is discussed above, the transition probability of a Brownian particle under the influence of the potential $U(x)$ is given by

$$T(x_2, t_2 | x_1, t_1) = N \int \exp \left\{ - (\zeta/2kT) J(\text{path}) \right\} d(\text{path}), \quad (3.12)$$

$$J = \int_{t_1}^{t_2} \mathcal{L}(\dot{x}, x, t) dt,$$

$$\mathcal{L} = 1/2 \cdot \{d\mathbf{x}/dt + (1/\zeta)\nabla U\}^2. \quad (3 \cdot 13)$$

The correspondence with the case of quantum mechanics developed in Appendix is

$$i/\hbar \rightarrow -(\zeta/2kT) \quad (= -1/2D) \quad \text{and} \quad L \rightarrow \mathcal{L}. \quad (3 \cdot 14)$$

Then we have from eq. (A·5) in Appendix

$$\pi_x = \dot{x} + (1/\zeta)\partial U/\partial x, \quad (3 \cdot 15)$$

$$\mathcal{H} = 1/2 \cdot \{d\mathbf{x}/dt + (1/\zeta)\nabla U\} \cdot \{d\mathbf{x}/dt - (1/\zeta)\nabla U\}, \quad (3 \cdot 16)^*$$

which yield immediately various relations: i.e., the differential representation of momentum is given from eq. (A·9) by

$$\pi_x = \dot{x} + (1/\zeta)\partial U/\partial x \rightarrow -2D(\partial/\partial x),$$

in a representation which makes x diagonal, and Fokker-Planck equation is given from eq. (A·10) by

$$\begin{aligned} 2D(\partial W/\partial t) &= \mathcal{H}(\pi, x)W \\ &= (1/2)\pi \cdot (\pi - (2/\zeta)\nabla U)W \\ &= 2D\nabla \cdot (D\nabla - (1/\zeta)\nabla U)W, \end{aligned} \quad (3 \cdot 18)$$

or

3·2 O-U process (I)

Putting $dx/dt = u$, the path integral formulation of a free Brownian particle is given by

$$W(u_2, t_2 | u_1, t_1) = N \int \exp \{ - (1/2\zeta kT) I \} d(\text{path}), \quad (3 \cdot 19)$$

$$I = 1/2 \cdot \int_{t_1}^{t_2} (m\dot{u} + \zeta u)^2 dt.$$

In this case the path in u -space suffices our present purpose and the specification of the path in x space is not necessarily required. This equation was given by Onsager-Machlup.⁵⁾ Momentum operator $\pi_x = m(m\dot{u}_x + \zeta u_x)$ is represented by a differential operator

$$\pi_x = m(m\dot{u}_x + \zeta u_x) \rightarrow -2\zeta kT(\partial/\partial u_x) \quad (3 \cdot 20)$$

and Fokker-Planck equation is given by

$$2\zeta kT(\partial W/\partial t) = \mathcal{H}(\pi, u)W \quad (3 \cdot 21)$$

* The physical meaning of the Hamiltonian operator in Brownian motion is the logarithmic decrement (see Fokker-Planck equation (2·2') (L_1 is Hamiltonian apart from constant factor) and (2·8'), and see also the final paragraph of the Appendix).

$$\begin{aligned}\mathcal{H}(\boldsymbol{\pi}, \mathbf{u}) &= (1/2) (\dot{m}\mathbf{u} + \boldsymbol{\zeta}\mathbf{u}) \cdot (\dot{m}\mathbf{u} - \boldsymbol{\zeta}\mathbf{u}) \\ &= (1/2m^2) \boldsymbol{\pi} \cdot (\boldsymbol{\pi} - 2\boldsymbol{\zeta}\mathbf{u})\end{aligned}\quad (3.22)$$

or,

$$\partial W / \partial t = \beta \nabla_{\mathbf{u}} \cdot (\mathbf{u}W) + (kT\beta/m) \Delta_{\mathbf{u}} W, \quad \beta = \zeta/m \quad (3.23)$$

which is the eq. (277) in Chandrasekhar's paper.¹¹⁾

3.3 O—U process (II)

In 3.2 the specification of positions in initial and final states was not required to the expression of transition probability, and only the path in \mathbf{u} -space is considered. However one may take into account the path in \mathbf{x} -space. General treatment of this case requires further development of dynamics of a system whose Lagrangian involves \mathbf{x} , \mathbf{u} and $\dot{\mathbf{u}}$. This generalization is easily carried out following the method of Ostrogradsky¹²⁾ and is outlined in § 4. However in our case of free particle we can derive Fokker-Planck equation without recourse to this generalization. The transition probability is represented by

$$\begin{aligned}T(\mathbf{x}_2, \mathbf{u}_2, t_2 | \mathbf{x}_1, \mathbf{u}_1, t_1) &= N \int \exp \{ - (1/2\zeta kT) I \} d(\text{path}) \\ I &= 1/2 \cdot \int_{t_1}^{t_2} (\dot{m}\mathbf{u} + \boldsymbol{\zeta}\mathbf{u})^2 dt.\end{aligned}\quad (3.24)$$

Lagrangian of this case is same as that of the preceding case, because the fluctuation force is same. The path of eq. (3.24), however, is to be taken in \mathbf{xu} -space, satisfying the condition $\mathbf{u} = d\mathbf{x}/dt$, which means that the arbitrary path in \mathbf{xu} -space cannot be followed. One may rather say that the path in \mathbf{x} -space is determined by the condition $\mathbf{x} = \int \mathbf{u} dt$. In consequence the Fokker-Planck equation is obtained by adding a term $\mathbf{u} \cdot \nabla_{\mathbf{x}} W$ to the left-hand side of the above equation (3.23) of O—U process (I). One obtains

$$\partial W / \partial t + \mathbf{u} \cdot \nabla_{\mathbf{x}} W = \beta \nabla_{\mathbf{u}} \cdot (\mathbf{u}W) + (kT\beta/m) \Delta_{\mathbf{u}} W. \quad (3.25)$$

This equation is already obtained by Chandrasekhar (eq. (249) in his paper¹¹⁾).

§ 4. “Heisenberg” representation of the theory of Brownian motion

In the present section we shall develop the operator formalism of the theory of Brownian motion, which is analogous to the “Heisenberg” representation in quantum mechanics.

4.1 Generalities

From now on we will use Dirac's notations, in which a state of Brownian motion is described by an abstract vector, the “ket” $|\rangle$ or the “bra” $\langle|$ say, and a physical quantity by an operator* acting on it.

* We shall denote an abstract operator by the under-bar, for instance, \underline{x} for the position \mathbf{x} or $\underline{\mathcal{H}}$ for the logarithmic decrement \mathcal{H} .

On the analogous way to quantum mechanics, assume that the set of eigenvectors of any operator be complete orthonormal one. However, in the theory of Brownian motion, it seems that this assumption may as well be required only for the position \underline{x} and the logarithmic decrement $\underline{\mathcal{L}}$.

Now, the transition probability density for the motion of the particle from $(x_1, t_1)^*$ to (x_2, t_2) is written down in the form

$$T(x_2, t_2 | x_1, t_1) = \sum_i u_i(x_2) e^{-(1/\gamma)\alpha_i(t_2-t_1)} v_i^*(x_1),$$

where γ stands for $2D$ in Wiener process or $2kT\zeta$ in O-U process and $u_i(x)$ and $v_i(x)$ are the representatives $\langle x | \alpha_i \rangle$ and $\langle \alpha_i | x \rangle$, respectively, of the eigenvector of $\underline{\mathcal{L}}$ belonging to the eigenvalue α_i in the diagonal representation of \underline{x} . In virtue of the condition for normalized completeness

$$\sum_i |\alpha_i\rangle \langle \alpha_i| = \underline{1},$$

the expression (4.1) is easily reduced to the compact form

$$T(x_2, t_2 | x_1, t_1) = \langle x_2 | \underline{T}(t_2, t_1) | x_1 \rangle \quad (4.2)$$

with

$$\underline{T}(t_2, t_1) = \exp[-(1/\gamma)(t_2 - t_1)\underline{\mathcal{L}}]. \quad (4.3)$$

Obviously this is the transformation operator governing the evolution of Brownian motion, obeying the Fokker-Planck equation in the operational form

$$\gamma(\partial/\partial t_2) \underline{T}(t_2, t_1) = -\underline{\mathcal{L}} \underline{T}(t_2, t_1), \quad (4.4)$$

with the initial condition

$$\underline{T}(t_1, t_1) = \underline{1}. \quad (4.5)$$

Moreover, it is evident that the operator \underline{T} satisfies Kolmogoroff-Chapman's equation

$$\underline{T}(t_2, t_1) = \underline{T}(t_2, t_3) \underline{T}(t_3, t_1). \quad (4.6)$$

The normalization condition is usually expressed as

$$\int_V T(x, t | x_0, t_0) dx = 1, \quad (4.7)$$

where V is volume of the total system. Thus, the normalization condition for the operator \underline{T} turns out to be

$$\langle V | \underline{T}(t, t_0) | x_0 \rangle = 1 \quad (4.8)$$

from (4.2) and (4.7), where $|V\rangle = \int_V dx |x\rangle$. Consequently, the condition (4.7) requires the property

* In O-U processes one must take other independent variables i.e., velocity u or x and u , however, for the sake of simplicity x is used throughout for such independent variables.

$$\langle V | \underline{\mathcal{H}} | x_0 \rangle = 0 \quad (4.9)$$

for the operator $\underline{\mathcal{H}}$. It is true that this requirement is of great importance as a fundamental character of the operator $\underline{\mathcal{H}}$ on a par with the Hermitian character of the Hamiltonian operator in quantum mechanics, but it is always possible to construct the operator $\underline{\mathcal{H}}$ with such a character in the framework of the theory as formulated in the preceding section. We shall see this fact later. Further it holds that

$$\langle V | \underline{T}(t, t_0) | f \rangle = 1,$$

if

$$|f\rangle = \int_V |x\rangle dx f(x) \quad \text{and} \quad \int_V f(x) dx = 1.$$

In what follows we shall treat only the operator $\underline{\mathcal{H}}$ with the character (4.9).

Next, let us consider the mean value of a quantity $A(x)$ after $(t-t_0)$ for an initial distribution $f(x)$ at t_0 , that is,

$$\langle A \rangle_{(t-t_0),f} \equiv \iint_V A(x) T(x, t | x_0, t_0) f(x_0) dx dx_0. \quad (4.10)$$

This easily rewritten in the operational form

$$\langle A \rangle_{(t-t_0),f} = \langle V | \underline{A}(\underline{x}(t)) | f \rangle, \quad (4.11)$$

where $\underline{x}(t)$ is the time-dependent operator

$$\underline{x}(t) = \exp[(1/\gamma) \underline{\mathcal{H}}(t-t_0)] \underline{x} \exp[-(1/\gamma) \underline{\mathcal{H}}(t-t_0)], \quad (4.12)$$

obeying the "Heisenberg" equation of motion

$$\gamma(d\underline{x}(t)/dt) = [\underline{\mathcal{H}}, \underline{x}(t)] \quad (4.13)$$

and the initial condition $\underline{x}(0) = \underline{x}$. The proof of the equality of (4.11) to (4.10) is performed as follows; since the operator $\underline{A}(\underline{x}(t))$ has the same time-dependence as (4.12) for $\underline{x}(t)$, the right-hand side of (4.11) becomes

$$\begin{aligned} \langle V | \underline{A}(\underline{x}(t)) | f \rangle &= \int_V \langle V | \underline{T}(t_0, t) | x \rangle dx \langle x | \underline{A}(\underline{x}) \underline{T}(t, t_0) | f \rangle \\ &= \int_V \langle x | \underline{A}(\underline{x}) \underline{T}(t, t_0) | f \rangle dx \\ &= \iint_V A(x) T(x, t | x_0, t_0) f(x_0) dx dx_0, \end{aligned}$$

where we have made use of (4.8), $\langle x | \underline{A}(\underline{x}) | x' \rangle = A(x) \delta(x-x')$ and $\langle x'' | f \rangle = f(x'')$. Thus the problem turns to solving the "Heisenberg" equation of motion (4.13) and then computing the matrix element (4.11). We shall show a few illustrations for some

simple problems in the following sub-sections.

4.2 Wiener process

As given in § 3, the "Lagrangian" for the Wiener process has the form

$$\mathcal{L} = (1/2) (\dot{\underline{x}} + (1/\zeta) \nabla U(\underline{x}))^2 \quad (4.14)$$

for the given external force with the potential function U , from which we get the operators

$$\left. \begin{aligned} \underline{\pi} &= \underline{x} + (1/\zeta) \nabla U(\underline{x}), \\ \mathcal{H} &= (1/2) \underline{\pi} \cdot (\underline{\pi} - (2/\zeta) \nabla U(\underline{x})). \end{aligned} \right\} \quad (4.15)$$

Here it is to be noted that the order of factors in \mathcal{H} has been chosen to satisfy the condition (4.9). The commutation relation between $\underline{\pi}$ and \underline{x} is obtained from (A.13):

$$[\pi_i, x_j] = -2D\delta_{ij}. \quad (4.16)$$

Hence their matrix elements become

$$\left. \begin{aligned} \langle x | \underline{\pi} | x_0 \rangle &= -2D \nabla \delta(\underline{x} - \underline{x}_0), \\ \langle x | \underline{x} | x_0 \rangle &= \underline{x}_0 \delta(\underline{x} - \underline{x}_0). \end{aligned} \right\} \quad (4.17)$$

which give us the relations

$$\langle V | \underline{x} | x_0 \rangle = \underline{x}_0, \quad \langle V | \underline{\pi} | x_0 \rangle = 0. \quad (4.18)$$

For simplicity, let us consider the motion in the case where there exists no external force, that is, $U=0$. The equation of motion (4.13) gives us the equation for free particle

$$d^2 \underline{x}(t)/dt^2 = 0 \quad (4.19)^*$$

and imposes the initial conditions on $\underline{x}(t)$ as follows;

$$\dot{\underline{x}}(0) = \underline{\pi} \quad (4.20)$$

besides the trivial one

$$\underline{x}(0) = \underline{x}.$$

Thus, we get the solutions

$$\underline{x}(t) = \underline{x} + \underline{\pi}t. \quad (4.21)$$

If we take the initial condition $f(\underline{x}) = \delta(\underline{x} - \underline{x}_0)$, the mean value of \underline{x} is easily computed from (4.21) and (4.18):

$$\langle \underline{x} \rangle_{t, x_0} = \underline{x}_0. \quad (4.22)$$

The mean square deviation from the initial position is similarly obtained:

* Note that $\dot{\underline{\pi}}(t) = 0$.

$$\begin{aligned} \langle (x_i - x_{0i})^2 \rangle_{t, x_0} &= \langle x_i^2 \rangle_{t, x_0} - x_{0i}^2 \\ &= \langle V | x_i^2 + \pi_i^2 t^2 + t (\pi_i x_i + x_i \pi_i) | x_0 \rangle - x_{0i}^2 = 2Dt, \end{aligned} \quad (4.23)$$

because of (4.18) and the relations

$$\begin{aligned} \langle V | \pi_i^2 | x_0 \rangle &= 0, \\ \langle V | \pi_i x_i + x_i \pi_i | x_0 \rangle &= 2D + \langle V | \pi_i x_i | x_0 \rangle, \\ \langle V | \pi_i x_i | x_0 \rangle &= x_{0i} \langle V | \pi_i | x_0 \rangle = 0. \end{aligned}$$

The transition probability is given by the symbolical form

$$T(\mathbf{x}, t | \mathbf{x}_0, 0) = \exp[2DtP^2] \delta(\mathbf{x} - \mathbf{x}_0). \quad (4.24)$$

Thus, we obtain the usual expression (2.4), by using $\delta(x) = (1/2\pi) \int_{-\infty}^{\infty} e^{ikx} dk$, for a system in a space with infinite extension. For the system in a cube (edge L) with the rigid wall, one has only to use the expression

$$\begin{aligned} \delta(\mathbf{x} - \mathbf{x}_0) &= (2/L)^3 \sum_{l=1}^{\infty} \sin(l\pi x_l/L) \sin(l\pi x_{0l}/L) \\ &\times \sum_{m=1}^{\infty} \sin(m\pi x_2/L) \sin(m\pi x_{02}/L) \sum_{n=1}^{\infty} \sin(n\pi x_3/L) \sin(n\pi x_{03}/L). \end{aligned}$$

4.3 O—U process

For simplicity, we shall discuss the O-U process only in case there exists no external force for the case of 3.3. The “Lagrangian” given in § 3 is

$$\mathcal{L} = (1/2) (m\dot{\mathbf{u}} + \boldsymbol{\zeta}\mathbf{u})^2, \quad \mathbf{u} = \dot{\mathbf{x}}$$

for this process. In order to deal with the velocity variables \mathbf{u} like independent ones and to construct the canonical formalism, it is more convenient to make use of the modified Lagrangian¹²⁾

$$\mathcal{L}' = (1/2) (m\dot{\mathbf{u}} + \boldsymbol{\zeta}\mathbf{u})^2 + \mathbf{p} \cdot (\dot{\mathbf{x}} - \mathbf{u}), \quad (4.25)$$

\mathbf{p} being the so-called “Lagrange’s undetermined coefficients.” The Lagrangian \mathcal{L}' generates the “momentums” $\boldsymbol{\pi}$ and \mathbf{p} canonically conjugate to the variables \mathbf{u} and \mathbf{x} , respectively, as follows;

$$\left. \begin{aligned} \pi_i &= \partial \mathcal{L}' / \partial \dot{u}_i = m(\dot{u}_i + \zeta u_i), \\ p_i &= \partial \mathcal{L}' / \partial \dot{x}_i = p_i. \end{aligned} \right\} \quad (4.26)$$

Hence, it turns out that the Lagrange’s coefficient \mathbf{p} is the “momentum” conjugate to the position \mathbf{x} . The operator \mathcal{H} has the form

$$\mathcal{H} = \mathbf{u} \cdot \mathbf{p} + (1/2m^2) \boldsymbol{\pi} \cdot (\boldsymbol{\pi} - 2m\boldsymbol{\zeta}\mathbf{u}). \quad (4.27)$$

in which the ordering of factors ensures the condition (4.9). From (A.13) one gets

$$\left. \begin{aligned} [\pi_i, u_j] &= -2kT\zeta\delta_{ij}, \\ [p_i, x_j] &= -2kT\zeta\delta_{ij}. \end{aligned} \right\} \quad (4.28)$$

The matrix elements of \underline{x} , \underline{u} , \underline{p} , $\underline{\pi}$ and $\underline{\mathcal{G}}$ are written down in the following form:

$$\left. \begin{aligned} \langle x, u | \underline{x} | x', u' \rangle &= x' \delta(x - x') \delta(u - u'), \\ \langle x, u | \underline{u} | x', u' \rangle &= u' \delta(x - x') \delta(u - u'), \\ \langle x, u | \underline{p} | x', u' \rangle &= -2kT\zeta \nabla_x \delta(x - x') \delta(u - u'), \\ \langle x, u | \underline{\pi} | x', u' \rangle &= -2kT\zeta \nabla_u \delta(x - x') \delta(u - u'), \\ \langle x, u | \underline{\mathcal{G}} | x', u' \rangle &= -2kT\zeta \{ \underline{u} \cdot \underline{\Gamma}_x - (kT\beta/m) \Gamma_u^2 - \beta \underline{\Gamma}_u \cdot \underline{u} \} \delta(x - x') \delta(u - u'). \end{aligned} \right\} \quad (4.29)$$

As the results of (4.29), it holds that

$$\left. \begin{aligned} \langle V_{xu} | \underline{x} | x_0, u_0 \rangle &= x_0, \quad \langle V_{xu} | \underline{u} | x_0, u_0 \rangle = u_0, \\ \langle V_{xu} | \underline{p} | x_0, u_0 \rangle &= \langle V_{xu} | \underline{\pi} | x_0, u_0 \rangle = 0, \\ \langle V_{xu} | \underline{x} \underline{\pi}_j | x_0, u_0 \rangle &= \langle V_{xu} | u_i p_j | x_0, u_0 \rangle = 0, \\ \langle V_{xu} | \underline{p}_i \underline{x}_j | x_0, u_0 \rangle &= \langle V_{xu} | \underline{\pi}_i u_j | x_0, u_0 \rangle = 0, \\ \langle V_{xu} | \underline{p}_i \underline{x}_i + \underline{x}_i \underline{p}_i | x_0, u_0 \rangle &= \langle V_{xu} | \underline{\pi}_i u_i + u_i \underline{\pi}_i | x_0, u_0 \rangle = 2kT\zeta. \end{aligned} \right\} \quad (4.30)$$

As in the Wiener process, the "Heisenberg" equation of motion gives us the equations

$$\left. \begin{aligned} (d^2 \underline{u}(t)/dt^2) - \beta^2 \underline{u}(t) &= -(1/m) \underline{p}, \\ \underline{u}(t) = d\underline{x}(t)/dt, \quad d\underline{p}(t)/dt &= 0 \end{aligned} \right\} \quad (4.31)$$

and the initial conditions

$$\left. \begin{aligned} \dot{\underline{x}}(0) &= \underline{u}, \quad \ddot{\underline{x}}(0) = \dot{\underline{u}}(0) = (1/m) \underline{\pi} - \beta \underline{u}, \\ \ddot{\underline{x}}(0) &= \ddot{\underline{u}}(0) = \beta^2 \underline{u} - (1/m) \underline{p}, \end{aligned} \right\} \quad (4.32)$$

besides the trivial one

$$\underline{x}(0) = \underline{x}.$$

The equations (4.31) under the initial conditions (4.32) are readily integrated with the result

$$\begin{aligned} \underline{x}(t) &= \underline{x} + \underline{u} (1 - e^{-\beta t}) / \beta + \underline{p} (1/m\beta^3) (\beta t - \sinh \beta t) \\ &\quad + \underline{\pi} (1/m\beta^3) (\cosh \beta t - 1), \end{aligned} \quad (4.33)$$

from which we obtain the mean value of the position

$$\langle \underline{x} \rangle_{t, (x_0, u_0)} = \underline{x}_0 + \underline{u}_0 (1 - e^{-\beta t}) / \beta, \quad (4.34)$$

and the mean square deviation from the initial position

$$\begin{aligned} \langle (x_i - x_{0i})^2 \rangle_{t, (x_0, u_0)} &= u_{0i}^2 (1 - e^{-\beta t})^2 / \beta^2 \\ &+ (kT/m\beta^2) [2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}]. \end{aligned} \quad (4.35)$$

In the derivation of (4.34) and (4.35), the relations (4.30) have been used.

It is easily shown that the symbolical solution for the transition probability

$$\begin{aligned} T(x, u, t | x_0, u_0, 0) \\ = \exp[-t \{ \mathbf{u} \cdot \nabla_x - (kT\beta/m) \nabla_u^2 - \beta \nabla_u \cdot \mathbf{u} \}] \delta(\mathbf{x} - \mathbf{x}_0) \delta(\mathbf{u} - \mathbf{u}_0) \end{aligned} \quad (4.36)$$

is reduced to the conventional expression (3.25) in a space with infinite extension, while it needs to use the special expression of the δ -functions in the case of the system with other particular boundary conditions.

§ 5. Quantum statistics

For a while we shall treat a one-particle system in a heat bath as an illustrative example of our formulation of quantum statistics. The extension to a many-particle system will be discussed later.

In virtue of the correspondence as mentioned in § 2, it is natural to define the density matrix for a one-particle system with the Lagrangian $L(x, \dot{x})$ by the path integral

$$\rho(x_2, \beta_2' | x_1, \beta_1') = N \int \exp \{ - (1/k) I(\text{path}) \} d(\text{path}) \quad (5.1)$$

with

$$I(\text{path}) = \int_{\beta_1'}^{\beta_2'} G(x(\beta'), \dot{x}(\beta')) d\beta',$$

where the function G is to be obtained from $L(x, \dot{x})$ by the replacement

$$G(x(s), \dot{x}(s)) = -L(x(s), (i/\hbar) k \dot{x}(s)). \quad (5.2)$$

For the conservative system with the Lagrangian $L = T - V$ (in which T and V are the kinetic and potential energies of the system, respectively), this relation becomes

$$G = (m/2) (k^2/\hbar^2) \dot{x}(s)^2 + V. \quad (5.3)$$

Here it is noted that, in quantum statistical mechanics, the dot must not stand for the derivative with respect to time but to inverse temperature. The conjugate momentum to the position x

$$\pi = \partial G / \partial \dot{x} = m (k^2/\hbar^2) \dot{x} \quad (5.4)$$

is considered to be the differential operator

$$\pi = -k(\partial/\partial x) \quad (5.5)$$

in the diagonal representation of the position operator \underline{x} . In such a representation, we get the "Hamiltonian" operator

$$\begin{aligned} H &= -(m/2) (\hbar^2/\hbar^2) \dot{x}^2 + V \\ &= -(\hbar^2/2m) (\partial^2/\partial x^2) + V. \end{aligned} \quad (5.6)$$

Thus the density matrix must obey the equation

$$-k(\partial\rho(x_2, \beta_2'|x_1, \beta_1')/\partial\beta_2') = [- (\hbar^2/2m \cdot \partial^2/\partial x_2^2) + V(x_2)] \rho(x_2, \beta_2'|x_1, \beta_1'), \quad (5.7)$$

being the well-known Bloch equation.

The abstract operators $\underline{\pi}$ and \underline{x} satisfy the commutation relation

$$[\underline{\pi}, \underline{x}] = -k \quad (5.8)$$

from (A.13). If we treat the variable $\beta = (1/kT)$, in the place of $\beta' = (1/T)$, like the time variable in quantum mechanics, the commutation relation becomes

$$[\underline{\pi}, \underline{x}] = -1. \quad (5.8)$$

Now, if we define the abstract operator $\underline{\rho}(\beta_2', \beta_1')$ by means of the matrix element

$$\langle x_2 | \underline{\rho}(\beta_2', \beta_1') | x_1 \rangle = \rho(x_2, \beta_2' | x_1, \beta_1'), \quad (5.9)$$

it must obey the operator equation

$$-k(\partial\underline{\rho}(\beta_2', \beta_1')/\partial\beta_2') = \underline{H}\underline{\rho}(\beta_2', \beta_1') \quad (5.10)$$

with the initial condition $\underline{\rho}(\beta_1', \beta_1') = 1$, where

$$\underline{H} = -(1/2m) (\hbar^2/k^2) \underline{\pi}^2 + \underline{V}(\underline{x}). \quad (5.11)$$

The equation (5.10) has the formal solution

$$\underline{\rho}(\beta_2', \beta_1') = \exp[-(1/k)(\beta_2' - \beta_1')\underline{H}], \quad (5.12)$$

whose matrix element is the conventional expression for the density matrix

$$\rho(x_2, \beta_2' | x_1, \beta_1') = \sum_i \phi_i(x_2) e^{-(1/k)E_i(\beta_2' - \beta_1')} \phi_i^*(x_1). \quad (5.13)$$

Here we have used $\phi_i(x) = \langle x | E_i \rangle$, $\underline{H}|E_i\rangle = E_i|E_i\rangle$ and the normalized completeness condition $\sum_i |E_i\rangle\langle E_i| = 1$ for the set $\{|E_i\rangle\}$. In some simple cases the potential function is linear or quadratic in \underline{x} and $\underline{\pi}$, one can easily solve the operator equation—the "Heisenberg" equation of motion—for the temperature dependent operators defined by the relation as (4.12). However, we do not enter them.

Next, we shall discuss a many-particle problem. Let the density matrix for an N -particle system be $\rho^{(N)}$. In case there occurs no increase or decrease in the number of particles, the operator $\rho^{(N)}$ is similarly obtained by

$$\underline{\rho}^{(N)}(\beta_2', \beta_1') = \exp[-(1/k)(\beta_2' - \beta_1')\underline{H}^{(N)}], \quad (5.14)$$

where $\underline{H}^{(N)}$ is the Hamiltonian operator of the N -particle system. However, if the system contacts with the source of particles, the number of particles is indefinite and it is necessary to define the grand density operator

$$\underline{P}(\beta_2', \beta_1') = \sum_N e^{(1/k)\beta_2'\mu N} \underline{\rho}^{(N)} e^{-(1/k)\beta_1'\mu N}, \quad (5.15)$$

where μ is the chemical potential of particles. In this case, it is convenient to use the method of description given by the second quantization, in which $|E_i^{(N)}\rangle$ turns to the simultaneous eigenvector of the Hamiltonian operator \underline{H} and the number operator \underline{N} (in the second quantized form) belonging to the eigenvalues E_i and N , respectively. By making use of the normalized completeness condition

$$\sum_N \sum_i |E_i^{(N)}\rangle \langle E_i^{(N)}| = \underline{1},$$

it is shown that (5.15) is reduced to the simple form

$$\underline{P}(\beta_2', \beta_1') = \exp[-(1/k)(\beta_2' - \beta_1')\underline{H}], \quad (5.16)$$

where

$$\underline{H}' = \underline{H} - \mu \underline{N}. \quad (5.17)$$

The operator \underline{H}' represents the total energy including the interaction energy between the system and the source. The form of \underline{H}' is concretely expressed in terms of the second quantized field operators $\underline{\psi}(x)$ and $\underline{\psi}^*(x)$ as follows:

$$\left. \begin{aligned} \underline{H} &= \int \underline{\psi}^*(x) [-(\hbar^2/2m)\nabla^2 + V(x)] \underline{\psi}(x) dx \\ &+ \iint \underline{\psi}^*(x) \underline{\psi}^*(x') V_{12}(x, x') \underline{\psi}(x) \underline{\psi}(x') dx dx', \\ \underline{N} &= \int \underline{\psi}^*(x) \underline{\psi}(x) dx \end{aligned} \right\} \quad (5.18)$$

with

$$\left. \begin{aligned} [\underline{\psi}(x), \underline{\psi}^*(x')]_{\mp} &= \delta(x - x'), \\ [\underline{\psi}(x), \underline{\psi}(x')]_{\mp} &= [\underline{\psi}^*(x), \underline{\psi}^*(x')]_{\mp} = 0, \end{aligned} \right\} \quad (5.19)$$

where $[A, B]_{\mp} = \underline{A}\underline{B} \pm \underline{B}\underline{A}$ and one must take minus or plus, according as the particles are bosons or fermions. $V(x)$ is the potential function for the external forces and $V_{12}(x, x')$ for the inter-particle forces. Moreover we can construct $|E_i^{(N)}\rangle$ as

$$|E_i^{(N)}\rangle = \frac{1}{\sqrt{N!}} \int \dots \int dx^{(1)} \dots dx^{(N)} \underline{\psi}^*(x^{(1)}) \dots \underline{\psi}^*(x^{(N)}) |0\rangle \phi_i(x^{(1)}, \dots, x^{(N)}), \quad (5.20)$$

in which the vector $|0\rangle$ represents the vacuum state, having the properties

$$\underline{\psi}|0\rangle = 0, \quad \underline{N}|0\rangle = 0 \quad \text{and} \quad \underline{H}|0\rangle = 0 \quad (5.21)$$

The temperature dependent operators $\underline{\psi}(x, \beta')$ and $\underline{\psi}^*(x, \beta')$ are defined as

$$\left. \begin{aligned} \underline{\psi}(x, \beta') &= e^{(1/k)\beta' \underline{H}'} \underline{\psi}(x) e^{-(1/k)\beta' \underline{H}'} \\ \underline{\psi}^*(x, \beta') &= e^{(1/k)\beta' \underline{H}'} \underline{\psi}^*(x) e^{-(1/k)\beta' \underline{H}'} \end{aligned} \right\} \quad (5.22)$$

which satisfy the "Heisenberg" equation of motion

$$k(\partial \underline{\psi}(x, \beta') / \partial \beta') = -[\underline{\psi}(x, \beta'), \underline{H}'] \quad (5.23)$$

By using (5.18) and (5.19), this equation is reduced to the Bloch equation with the source terms as an operator equation. If the quantity

$$\rho^{(1)}(x_2, \beta_2' | x_1, \beta_1') = \langle 0 | \underline{\psi}(x_2, \beta_2') \underline{\psi}^*(x_1, \beta_1') | 0 \rangle \quad (5.24)$$

is defined by the above operators, it obeys the Bloch equation (5.7) and, by making use of (5.21), it becomes

$$\rho^{(1)}(x_2, \beta_2' | x_1, \beta_1') = \sum_i \langle 0 | \underline{\psi}(x_2) | E_i^{(1)} \rangle e^{-(1/k)E_i(\beta_2' - \beta_1')} \langle E_i^{(1)} | \underline{\psi}^*(x_1) | 0 \rangle. \quad (5.25)$$

Here we have used the fact that $\underline{\psi}^*|0\rangle$ is an eigenvector of the operator N belonging to the eigenvalue 1, that is, $\langle 0 | \underline{\psi} | E_i^{(N)} \rangle = 0$ and $\langle E_i^{(N)} | \underline{\psi}^* | 0 \rangle = 0$ for $N \neq 1$. Since the function $\langle 0 | \underline{\psi}(x_2) | E_i \rangle$ can be interpreted as the amplitude for finding one particle at x in the state $|E_i\rangle$, the expression (5.25) is just the conventional one for the one-particle density matrix. For the two-particle problem, one can define

$$\begin{aligned} \rho^{(2)}(x_2^{(1)}, x_2^{(2)}, \beta_2' | x_1^{(1)}, x_1^{(2)}, \beta_1') \\ = \langle 0 | \underline{\psi}(x_2^{(1)}, \beta_2') \underline{\psi}(x_2^{(2)}, \beta_2') \underline{\psi}^*(x_1^{(1)}, \beta_1') \underline{\psi}^*(x_1^{(2)}, \beta_1') | 0 \rangle \end{aligned}$$

and, furthermore, similar definitions of the density matrices for the many-particle problems are possible. Such a formulation^{(1), (14)} has the complete analogy for the theory of the propagators^(1,3) in the quantized theory of field and then various field-theoretic techniques may be powerful for calculation of the density matrix in case of more general sources of particle.

Finally it may be interesting to note that the quantity $\rho^{(1)}$ can be expressed by the functional integral—not to be the sum over paths, but to be sum over fields—

$$\begin{aligned} \rho^{(1)}(x_2, \beta_2' | x_1, \beta_1') &= N \int \underline{\psi}(x_2, \beta_2') \underline{\psi}^*(x_1, \beta_1') \\ &\times \exp \left[- (1/k) \int_{\beta_1'}^{\beta_2'} \int_{-\infty}^{\infty} dx G(\underline{\psi}(x, \beta), \partial \underline{\psi}(x, \beta) / \partial x, \partial \underline{\psi}(x, \beta) / \partial \beta) \right] \delta \underline{\psi} \delta \underline{\psi}^*, \end{aligned} \quad (5.26)$$

where the function G is the field-theoretic Lagrangian which will give us the Hamiltonian like (5.17) with (5.18).

§ 6. Conclusions

An essential feature which distinguishes quantum mechanics from classical mechanics is the commutation relation

$$[p, q] = \hbar/i. \quad (6.1)$$

In the limit $\hbar \rightarrow 0$, we are led to classical mechanics.

In the above treatment, we have various commutation relations, for example, for a Wiener process discussed in §§ 3.1 and 4.2

$$[\pi_i, x_j] = -2D\delta_{ij}, \quad (6.2)$$

for an O-U process discussed in §§ 3.2, 3.3 and 4.3

$$\begin{aligned} [\pi_i, u_j] &= -2kT\zeta\delta_{ij}, \\ [p_i, x_j] &= -2kT\zeta\delta_{ij} \end{aligned} \quad (6.3)$$

and for quantum statistics

$$[\pi, x] = -k. \quad (6.4)$$

In these cases, in the limit $D \rightarrow 0$, $kT\zeta \rightarrow 0$, or $k \rightarrow 0$ every dynamical quantities commute each other and deterministic description is possible, corresponding to the case of classical limit of quantum mechanics.

In quantum mechanics of a free particle the transformation function is given by

$$K(x, t|x_0, t_0) = [-im/2\pi\hbar(t-t_0)]^{1/2} \exp[im(x-x_0)^2/2\hbar(t-t_0)]$$

in one dimensional case. This is the same form as eq. (2.4) for Wiener process, with the correspondence $im/\hbar \rightarrow 1/2D$. This is the reason why the commutation relation (4.16) holds for Brownian motion. Inversely quantum mechanics may be said as a Wiener process in the sense that in q -representation the velocity cannot be determined. However the important difference between quantum mechanics and Brownian motion lies in the structure of the Lagrangians which appear in path integral formulations. In other words quantum mechanics cannot be described by introducing a fluctuation force as is done in Langevin's equation of Brownian motion. It might be possible to construct a modified quantum mechanics by means of introducing higher time-derivatives of velocity in Lagrangian in a way analogous to the relationship between O-U process and Wiener process, so as to make possible a simultaneous determination of position and velocity of a particle. However such a direct modification of quantum mechanics would be of less significance, because the Hamiltonian in this formulation would not be afforded any physical meaning.

Appendix. Feynman's and Schwinger's formalism of quantum mechanics

For the convenience of the applications given in the present paper, we shall briefly interpret two new formalisms of quantum mechanics obtained by Feynman and Schwinger, in which the theory is constructed directly from the Lagrangian without resort to the canonical formalism. In particular, the latter gives us the unified description for the separate fundamental requirements of the conventional theory, that is, the commutation

relations and the equation of motion, etc.

We first mention Feynman's theory, on the basis of which one explain the dynamical principle in Schwinger's theory.* Let the Lagrangian, L , of the system be a function of the independent dynamical variables, α_i 's, and their time derivatives, $\dot{\alpha}_i$'s. For simplicity, we neglect the dependence of L on the higher time derivatives. The extension to the system with the Lagrangian containing the higher derivatives is readily performed by the method given by Ostrogradsky⁽¹²⁾; for example, see the O-U process (II) in § 4. The quantization in Feynman's theory is attributed to the postulation that the transition amplitude for finding the values α_i'' 's for the dynamical quantities α_i 's at the time t_2 in the system with the initial state having the values α_i' 's for them at the time t_1 be obtained by evaluating a path integral as follows;

$$\langle \alpha'', t_2 | \alpha', t_1 \rangle = N \int \exp \left[i/\hbar \int_{t_1}^{t_2} L(\alpha(t), \dot{\alpha}(t)) dt \right] d(\text{path})_\alpha, \quad (\text{A} \cdot 1)$$

where N is a normalization constant. The matrix element of any dynamical quantity F between the above-mentioned states is defined by

$$\langle \alpha'', t_2 | F | \alpha', t_1 \rangle = N \int F(\text{path}) \exp \left[i/\hbar \int_{t_1}^{t_2} L(\alpha(t), \dot{\alpha}(t)) dt \right] d(\text{path})_\alpha \quad (\text{A} \cdot 2)$$

where F is an operator having (A·2) as its matrix element, corresponding to the classical dynamical quantity F . As is easily seen, Feynman's theory consists of evaluations of the c -number path integrals rather than the operator calculus. However, the operator defined by (A·2) has the same character as the corresponding one in the conventional theory and obeys Heisenberg equation of motion, in case the canonical variables can be considered. This means the equivalence of Feynman's formalism using the path integral to the conventional one containing operator. In order to perform this transformation of the description from Feynman's to the operator formalism in general cases, we must investigate the law of variation of the transition amplitude for the following three infinitesimal variations; i) to modify the system, $L \rightarrow L + \delta L$, ii) to displace the path, $\alpha(t) \rightarrow \alpha(t) + \delta^* \alpha(t)$, and iii) to change the interval of motion, $(t_2, t_1) \rightarrow (t_2 + \delta t_2, t_1 + \delta t_1)$. For such variations, the action function on a path

$$I_H = \int_{t_1}^{t_2} L(\alpha_H(t), \dot{\alpha}_H(t)) dt \quad (\text{A} \cdot 3) **$$

suffers the following variation

* Such an explanation of Schwinger's principle is offered by Dyson (*Advanced Quantum Mechanics*, Cornell University, (1951) p. 36). Many parts of the present section in this paper owe to the lectures given by Professor R. Utiyama, to whom the authors express their thanks. Readers may also refer to his work on quantum theory of the non-local field, Imamura, Sunakawa and Utiyama, *Prog. Theor. Phys.*, **11** (1954) 29

** The suffix H is used to denote a path or values on a path.

$$\begin{aligned}
\delta I_H &= \int_{t_1+\delta t_1}^{t_2+\delta t_2} \{L(\alpha_H + \delta^* \alpha_H, \dot{\alpha}_H + \delta^* \dot{\alpha}_H) + \delta_L L(\alpha_H, \dot{\alpha}_H)\} dt \\
&\quad - \int_{t_1}^{t_2} L(\alpha_H, \dot{\alpha}_H) dt \\
&= [\sum_i \pi_i(t_2) \delta \alpha_i'' - H(t_2) \delta t_2] - [\sum_i \pi_i(t_1) \delta \alpha_i' - H(t_1) \delta t_1] \\
&\quad + \int_{t_1}^{t_2} \{\delta_L L(\alpha_H, \dot{\alpha}_H) + \sum_i [L]_{\alpha_i H} \delta^* \alpha_{iH}\} dt, \tag{A.4}
\end{aligned}$$

where

$$\left. \begin{aligned} \pi_i &= \partial L / \partial \dot{\alpha}_i, \quad \delta \alpha_i = \delta^* \alpha_i + \dot{\alpha}_i \delta t, \\ H &= \sum_i \pi_i \dot{\alpha}_i - L, \\ [L]_{\alpha_i} &= \partial L / \partial \alpha_i - (d/dt) \partial L / \partial \dot{\alpha}_i. \end{aligned} \right\} \tag{A.5}$$

The variation of the amplitude is given as

$$\delta \langle \alpha'', t_2 | \alpha', t_1 \rangle = \langle \alpha'' + \delta \alpha'', t_2 + \delta t_2 | \alpha' + \delta \alpha', t_1 + \delta t_1 \rangle - \langle \alpha'', t_2 | \alpha', t_1 \rangle.$$

Thus we obtain from (A.1) and (A.2) the variational equation of the transition amplitude

$$\delta \langle \alpha'', t_2 | \alpha', t_1 \rangle = (i/\hbar) \langle \alpha'', t_2 | \delta \underline{I}(\alpha) | \alpha', t_1 \rangle. \tag{A.6}$$

Here $\delta \underline{I}(\alpha)$ is an operator composed of the c -number function (A.4) by means of the rule (A.2). Consequently, such a change in the description results in the appearance of the operators, instead of making the explicit use of the path integrals.

Now one may consider the equation (A.6) as a quantization condition, although it has been derived from (A.1). In fact, to postulate the equation (A.6) as a starting principle of quantization is equivalent to (A.1), because (A.1) can be derived from (A.6) as is easily proved. This is the *dynamical principle* required by Schwinger.

In Schwinger's theory, the kinematical postulations of quantum mechanics are just same as the conventional ones as follows. A state of the system of described by a vector in an abstract vector space, and a dynamical quantity is represented by a hermitian operator mapping a vector to another vector in the space. The transition amplitude for the process $(\alpha', t_1) \rightarrow (\alpha'', t_2)$ is given by the inner product, $\langle \alpha'', t_2 | \alpha', t_1 \rangle$, composed is the eigenvectors $|\alpha'', t_2\rangle$ and $|\alpha', t_1\rangle$ of the operator $\underline{\alpha}(t)$ belonging to the eigenvalue α'' at t_2 and α' at t_1 , respectively. An operator $\underline{F}(\underline{\alpha}(t))$ can be represented by the matrix element $\langle \alpha'', t_2 | \underline{F}(\underline{\alpha}(t)) | \alpha', t_1 \rangle$, which is an inner product composed of the vectors $|\alpha'', t_2\rangle$ and $\underline{F}(\underline{\alpha}(t)) |\alpha', t_1\rangle$. The set of the vector $|\alpha', t\rangle$ is a complete orthonormal one at any instant. Consequently, one gets

$$\langle \alpha'', t | \underline{F}(\underline{\alpha}(t)) | \alpha', t \rangle = F(\alpha') \delta(\alpha'' - \alpha') \tag{A.7}$$

and the multiplicative law for the matrix of \underline{F}

$$\langle \alpha'', t_2 | \underline{F} \underline{G} | \alpha', t_1 \rangle = \int \langle \alpha'', t_2 | \underline{F} | \alpha''', t_3 \rangle \langle \alpha''', t_3 | \underline{G} | \alpha', t_1 \rangle d\alpha'''. \quad (\text{A} \cdot 7')$$

From (A·7) and (A·7'), it holds that

$$\left. \begin{aligned} \langle \alpha'', t_2 | \underline{F}(\alpha(t)) | \alpha', t_1 \rangle &= \int \langle \alpha'', t_2 | \alpha''', t \rangle F(\alpha''') \langle \alpha''', t | \alpha', t_1 \rangle d\alpha''' \\ \langle \alpha'', t_2 | \alpha', t_1 \rangle &= \int \langle \alpha'', t_2 | \alpha''', t_3 \rangle \langle \alpha''', t_3 | \alpha', t_1 \rangle d\alpha''' \end{aligned} \right\} \quad (\text{A} \cdot 7'')$$

In addition to these kinematical postulations, the equation (4·6) is taken up as a dynamical postulation.

The dynamical principle (A·6) contains the following facts:

1) *Equation of motion.* For the special variations, $\delta\alpha'' = \delta\alpha' = 0$, $\delta t_2 = \delta t_1 = 0$ and $\delta_L L = 0$, it is expected that (A·6) becomes the variational expression of the equation of motion. Indeed, since the left-hand side of (A·6) vanishes for these variations, we obtain

$$0 = \langle \alpha'', t_2 | \int_{t_1}^{t_2} [L]_{\alpha_i} \delta^* \alpha_i dt | \alpha', t_1 \rangle,$$

from which the equation of motion written in terms of operators is derived in the classical form

$$[\underline{L}]_{\alpha_i} = 0. \quad (\text{A} \cdot 8)$$

In what follows, it is to be remembered that (A·8) always holds in so far as we investigate the real dynamical system.

2) *Representation of momentum operator by differential operator.* For the special variations, $\delta t_2 = \delta t_1 = 0$, $\delta\alpha' = 0$, $\delta_L L = 0$, $\delta\alpha_i'' = 0$ ($i \neq k$) but $\delta\alpha_k'' \neq 0$, (A·6) becomes

$$\begin{aligned} (\partial \langle \alpha'', t_2 | \alpha', t_1 \rangle / \partial \alpha_k'') \delta \alpha_k'' &= (i/\hbar) \langle \alpha'', t_2 | \underline{\pi}_k(t) | \alpha', t_1 \rangle \delta \alpha_k'' \\ &= (i/\hbar) \delta \alpha_k'' \int \langle \alpha'', t_2 | \underline{\pi}_k(t) | \alpha''', t \rangle \langle \alpha''', t | \alpha', t_1 \rangle d\alpha''', \end{aligned}$$

from which we obtain

$$\langle \alpha'', t | \underline{\pi}_k(t) | \alpha', t \rangle = (\hbar/i) (\partial / \partial \alpha_k'') \delta(\alpha'' - \alpha'). \quad (\text{A} \cdot 9)$$

3) *Schrödinger equation.* If we take all vanishing variations except $\delta t_2 \neq 0$ and put $\langle \alpha'', t_2 | \alpha', t_1 \rangle = \psi(\alpha'', t_2)$, (A·6) becomes

$$\begin{aligned} (d\psi(\alpha'', t_2) / dt_2) \delta t_2 &= - (i/\hbar) \langle \alpha'', t_2 | \underline{H}(\underline{\pi}(t_2), \underline{\alpha}(t_2)) | \alpha', t_1 \rangle \delta t_2 \\ &= - (i/\hbar) \delta t_2 \int \langle \alpha'', t_2 | \underline{H}(\underline{\pi}(t_2), \underline{\alpha}(t_2)) | \alpha''', t_2 \rangle \psi(\alpha''', t_2) d\alpha''', \end{aligned}$$

which results in Schrödinger equation

$$i\hbar (\partial \psi(\alpha', t) / \partial t) = H(\hbar/i \cdot \partial / \partial \alpha', \alpha') \psi(\alpha', t), \quad (\text{A} \cdot 10)$$

where (A·7) and (A·7') have been used.

The operator form of Schwinger's dynamical principle is also obtained. Now, let us consider the expression

$$\langle \alpha'', t_2 | \underline{F}(\alpha'(t_2)) | \alpha', t_2 \rangle = F(\alpha'') \delta(\alpha'' - \alpha')$$

and take its variation under the three infinitesimal variations as mentioned above. Since the right-hand side of it suffers the only variation $\alpha'' \rightarrow \alpha'' + \delta\alpha''$, the variation of the quantity $\langle \alpha'', t_2 | \underline{F} - \delta \underline{F} | \alpha', t_2 \rangle$, in which $\delta \underline{F}$ is an operator with the property $\langle \alpha'', t_2 | \delta \underline{F} | \alpha', t_2 \rangle = (F(\alpha'' + \delta\alpha'') - F(\alpha'')) \delta(\alpha'' - \alpha')$, vanishes, that is,

$$\begin{aligned} & \iint \{ \langle \alpha'' + \delta\alpha'', t_2 + \delta t_2 | \alpha''', t_3 \rangle \langle \alpha''', t_3 | \underline{F}(t_2 + \delta t_2) + \delta_L \underline{F}(t_2) - \delta \underline{F}(t_2) | \alpha^{IV}, t_3 \rangle \\ & \times \langle \alpha^{IV}, t_3 | \alpha' + \delta\alpha', t_2 + \delta t_2 \rangle \\ & - \langle \alpha'', t_2 | \alpha''', t_3 \rangle \langle \alpha''', t_3 | \underline{F}(t_2) | \alpha^{IV}, t_3 \rangle \langle \alpha^{IV}, t_3 | \alpha', t_3 \rangle \} d\alpha''' d\alpha^{IV} = 0. \end{aligned}$$

By using (A·6), this becomes

$$\begin{aligned} & (i/\hbar) \langle \alpha'', t_2 | [\delta \underline{I}(t_2), \underline{F}(t_2)] | \alpha', t_2 \rangle \\ & + \langle \alpha'', t_2 | \underline{F}(t_2 + \delta t_2) + \delta_L \underline{F}(t_2) - \delta \underline{F}(t_2) - \underline{F}(t_2) | \alpha', t_2 \rangle = 0. \end{aligned}$$

Thus we obtain the operator form of Schwinger's principle

$$\delta \underline{F}(t) - \dot{\underline{F}} \delta t - \delta_L \underline{F}(t) = (i/\hbar) [\delta \underline{I}(t), \underline{F}(t)]. \quad (\text{A} \cdot 11)$$

Here it is to be noted that, since (A·8) has been used for the one-side variation in the above derivation of (A·11), the operator $\delta \underline{I}(t)$ must be expressed as

$$\delta \underline{I}(t) = \sum_i \underline{\pi}_i(t) \delta \alpha_i - \underline{H}(t) \delta t + \int_{t_1}^{t_2} \delta_L L dt. \quad (\text{A} \cdot 12)$$

If one adds $G = - \sum_i \underline{\pi}_i \alpha_i$ to the Lagrangian from the outset—this addition is a canonical transformation and then does not modify the system (because the equation of motion is invariant)—, the operator $\delta \underline{I}(t)$ is changed as follows;

$$\delta \underline{I}(t) = - \sum_i \delta \underline{\pi}_i \alpha_i(t) - \underline{H}(t) \delta t + \int_{t_1}^{t_2} \delta_L L dt. \quad (\text{A} \cdot 12')$$

From (A·11) one can get the various operator equations:

1) *Commutation relations.* Let F be α_i . For the special variations $\delta t = 0$ and $\delta_L L = 0$, (A·11) becomes

$$-i\hbar \delta \alpha_i(t) = \sum_j \delta \alpha_j(t) [\underline{\pi}_j(t), \alpha_i(t)],$$

that is,

$$[\underline{\pi}_j(t), \alpha_i(t)] = (\hbar/i) \delta_{ij}. \quad (\text{A} \cdot 13)$$

For $F = \pi_i$, $\delta \pi_i = 0$ because π is independent of α . Thus it holds that

$$[\pi_j(t), \pi_i(t)] = 0. \quad (\text{A} \cdot 13')$$

If we use the expression (A·12'), we obtain

$$[\alpha_j(t), \alpha_i(t)] = 0. \quad (\text{A} \cdot 13'')$$

These are the canonical commutation relations.

2) *Heisenberg equation of motion.* For all vanishing variations except $\delta t \neq 0$, (A·11) gives us the equation

$$\dot{F}(t) = (i/\hbar) [H(t), F(t)]. \quad (\text{A} \cdot 14)$$

This is well-known Heisenberg equation of motion, which means that the operator \underline{H} is a generating operator for time translation. Putting all variations to zero except $\delta\alpha_k \neq 0$ in (A·11), we obtain

$$\partial F(\alpha) / \partial \alpha_k = (i/\hbar) [\pi_k, F(\alpha)]. \quad (\text{A} \cdot 15)$$

This is a general expression of the fact that the momentum operator and the angular momentum operator are generating operators for space translation and space rotation, respectively. Moreover, one can obtain from (A·11) and (A·12')

$$\partial F(\pi) / \partial \pi_k = -(i/\hbar) [\alpha_k, F(\pi)]. \quad (\text{A} \cdot 15')$$

Finally, we consider the conservation law of the dynamical quantity. If the action function is invariant under a transformation group, it holds that $\delta I = 0$ for an infinitesimal transformation belonging to this transformation group. Consequently, the expression (A·4) as an operator equation becomes

$$[\sum_i \pi_i \delta \alpha_i - H \delta t]_{t_2} = [\sum_i \pi_i \delta \alpha_i - H \delta t]_{t_1}, \quad (\text{A} \cdot 16)$$

where we have used $\delta_t L = 0$ and $[L]_{\alpha_k} = 0$. If we take a time translation $t \rightarrow t + \delta t$ for the invariant system under it, we obtain the conservation law of energy because of the fact $\delta \alpha = 0$ and $\delta t_2 = \delta t_1 = \delta t$ in this case. Similarly, for the invariant system under space displacement or rotation, (A·16) results in the conservation law of momentum or angular momentum. Thus the expression (A·16) is considered to be a unified representation of the conservation laws of various quantities.

References

- 1) E. Schrödinger, Sitzungsab. Preuss. Akad. Wiss. (1931), 144.
- 2) J. Metadier, Comptes Rendus 193 (1931), 1173.
- 3) R. Fürth, Zeits. Phys. 81 (1933), 143.
- 4) N. Wiener, Acta Math. 55 (1930), 117.
- 5) L. Onsager and S. Machlup, Phys. Rev. 91 (1953), 1505.
- 6) N. Hashitsume, Proc. Intern. Conf. Theor. Phys. (1953) Kyoto and Tokyo, 495.
- 7) R. P. Feynman, Rev. Mod. Phys. 20 (1948), 367.
- 8) E. Montroll, Comm. Pure & Appl. Math. 5 (1952), 415.
- 9) F. Bloch, Zeits. Phys. 74 (1932), 295.
- 10) R. P. Feynman, Phys. Rev. 91 (1953), 1291.
- 11) S. Chandrasekhar, Rev. Mod. Phys. 15 (1943), 1.
- 11a) N. Hashitsume, Busseiron Kenkyû 80 (1955), 52 (in Japanese).
- 12) E. Whittaker, Analytical Dynamics, 4th edition, p. 265.
- 13) H. Ichimura, Prog. Theor. Phys. 11 (1954), 374, 385, 519.
- 14) T. Matsubara, Prog. Theor. Phys. 14 (1955), 351.
- 15) P. T. Matthews and A. Salam, Nuovo Cimento, X, 2 (1955), 120.

Progress of Theoretical Physics, Vol. 16, No. 2, August 1956

Nucleon Magnetic Moments in ps - ps Intermediate Coupling Meson Theory

Akira KANAZAWA* and Masao SUGAWARA**†

*Physics Department, Hokkaido University, Sapporo

**Mathematical Physics, Royal Institute of Technology, Stockholm, Sweden†

(Received October 24, 1955)

Tomonaga intermediate coupling meson theory is applied to the symmetrical ps - ps meson theory, where we have taken account of the nucleon recoil, though non-relativistically, and the S -wave mesons, so as to give the pair damping automatically, besides the conventional P -wave mesons. Magnetic moments have been calculated, retaining up to three bound mesons and allowing virtual transitions to some excited states accompanied by single unbound meson emission. The results turn out to be too small to be consistent with the empirical values for reasonable cut-off momenta and coupling constants. Especially the effect of excited states has been shown to be negligible. Sachs' theorem has been obtained as regards the bound meson contribution.

§ 1. Introduction

Even if we confine our attention within the magnetic moments of nucleons, it is known that the present meson field theory is incapable in getting the quantitative agreement with the empirical values. To see whether the present meson theory itself is basically insufficient or the present method of calculation is rather unsatisfactory, we must carry out many more refined calculations than currently used. In order to give such an example, we have applied Tomonaga's intermediate coupling meson theory to the symmetrical ps - ps meson theory, taking account of some effects which have been neglected thus far, to calculate nucleon magnetic moments.

Heretofore¹⁾, the intermediate coupling theory was applied only to the case where we neglect the nucleon motion and retain only the derivative coupling term, and the effect of excited states has been completely neglected. In order to take account of these neglected effects at least partially, we have retained the non-relativistic recoil term of nucleon motion and some (nonlinear) term containing S -wave mesons, besides the derivative coupling term (P -wave meson term), after we have applied a canonical transformation²⁾³⁾ to the original Hamiltonian of ps - ps meson theory. We have considered, moreover, the virtual transitions to some excited states, accompanied by single unbound meson emission.

Our approximate Hamiltonian is constructed in section 2. As a first step of solving

†) On leave of absence from Physics Department, Hokkaido University.

††) Now, at Department of Physics, Purdue University, Lafayette, Indiana, U.S.A.

the resulting problem, we first eliminate the nucleon coordinate retained¹⁾ (section 3) and expand the meson field in terms of spherical waves (section 4). Then the Tomonaga approximation⁵⁾ is introduced in section 5 for both *S*- and *P*-wave mesons. The results of the eigenvalue solutions of the bound meson Hamiltonian are given in section 6. The calculated values of magnetic moments without and with the unbound meson effect (the effect of the excited states) are given and discussed in section 7. Main conclusions are summarized in the final section.

§ 2. Construction of approximate Hamiltonian

The Hamiltonian for coupled nucleon and meson fields, assuming symmetrical *ps-ps* theory, is given, in an external electromagnetic field A , by

$$H = \int \left\{ \psi^* \left[-i\alpha \left(\frac{\partial}{\partial \mathbf{x}} - ie \frac{1+\tau_3}{2} A \right) + \kappa \beta \right] \psi + \frac{1}{2} \sum_{\alpha} [\pi_{\alpha}^2 + \text{grad}^2 \phi_{\alpha} + \mu^2 \phi_{\alpha}^2] \right. \\ \left. + \sum_{\alpha} \psi^* i f \beta \gamma_5 \tau_{\alpha} \phi_{\alpha} \psi + e A \left(\phi_1 \frac{\partial}{\partial \mathbf{x}} \phi_2 - \phi_2 \frac{\partial}{\partial \mathbf{x}} \phi_1 \right) \right\} d\mathbf{x}, \quad (1)$$

using the ordinary notations.

To get an expression which is more suitable for our purpose, we first apply the Foldy transformation²⁾;

$$\exp(iS), \quad S \equiv \int \psi^* S \psi d\mathbf{x}, \quad S \equiv (1/2) \tan^{-1} \left(\sum_{\alpha} f \gamma_5 \tau_{\alpha} \phi_{\alpha} / \kappa \right), \quad (2)$$

and retain only the following terms

$$H' = \exp(iS) H \exp(-iS) \\ = \int \left\{ \psi^* \left[-i\alpha \left(\frac{\partial}{\partial \mathbf{x}} - ie \frac{1+\tau_3}{2} A \right) + \kappa \beta \right] \psi + \frac{1}{2} \sum_{\alpha} [\pi_{\alpha}^2 + \text{grad}^2 \phi_{\alpha} + \mu^2 \phi_{\alpha}^2] \right. \\ \left. + \psi^* \left[\sum_{\alpha} \frac{f}{2\kappa} \left(\frac{\kappa}{\kappa^*} \right)^2 \sigma \frac{\partial}{\partial \mathbf{x}} \tau_{\alpha} \phi_{\alpha} + (\kappa^* - \kappa) \beta \right] \psi \right. \\ \left. + e A \left[\psi^* \frac{f}{2\kappa} \left(\frac{\kappa}{\kappa^*} \right) \sigma (\phi_1 \tau_2 - \phi_2 \tau_1) \psi \right. \right. \\ \left. \left. + \left(\phi_1 \frac{\partial}{\partial \mathbf{x}} \phi_2 - \phi_2 \frac{\partial}{\partial \mathbf{x}} \phi_1 \right) \right] \right\} d\mathbf{x}, \quad (3)$$

where

$$\kappa^* = \kappa \left[1 + \sum_{\alpha} f^2 \phi_{\alpha}^2 / \kappa^2 \right]^{1/2}.$$

In (3), we have neglected other higher order terms* in $f/2\kappa$ and also a direct term (a

*) We have carried out the detailed calculations including some of these neglected terms. According to them, these effects are never quite negligible, nor large enough to modify the conclusions of this paper completely.

biquadratic term in ψ and ψ^*), since the one nucleon part of the latter can be shown to be a pure constant (infinite in general) except the higher order terms than those given above. Also we did not give explicitly in (3) those terms which contain the Dirac's odd matrixes, since they will be neglected in the next step. It is essential here that we have never expanded the nonlinear operator κ^* , which contains many higher order terms.

As we can take account of the nucleon motion only in the nonrelativistic approximation, we further apply the Foldy and Wouthuysen transformation³⁾ upon (3). After that, we come back to the configuration space representation as regards the nucleon motion. In the one-nucleon problem, we get finally the following approximate Hamiltonian

$$H'' = H''' + H''', \quad (5)$$

where

$$H''' = -\frac{1}{2\kappa}A + \frac{1}{2}\sum_{\alpha}\int[\pi_{\alpha}^2 + \text{grad}^2\phi_{\alpha} + \mu^2\phi_{\alpha}^2]d\mathbf{x} \\ + (\kappa^* - \kappa) + \sum_{\alpha}\frac{f}{2\kappa}\left(\frac{\kappa}{\kappa^*}\right)^2\sigma\frac{\partial}{\partial\mathbf{x}}\tau_{\alpha}\phi_{\alpha} \quad (6)$$

and

$$H'''' = -\frac{ef}{2\kappa}\left(\frac{\kappa}{\kappa^*}\right)\sigma A(\phi_1\tau_2 - \phi_2\tau_1) + \int eA\left[\phi_1\frac{\partial}{\partial\mathbf{x}}\phi_2 - \phi_2\frac{\partial}{\partial\mathbf{x}}\phi_1\right]d\mathbf{x} \\ + \frac{ie}{\kappa}\frac{1+\tau_3}{2}A\frac{\partial}{\partial\mathbf{x}} - \frac{e}{2\kappa}\frac{1+\tau_3}{2}\sigma\text{rot}A \\ - \frac{ef}{8\kappa^2}\sigma A\left[\left(\frac{\kappa}{\kappa^*}\right)^2\left(\sum_{\alpha}\tau_{\alpha}\pi_{\alpha} + \pi_3\right) + \left(\sum_{\alpha}\tau_{\alpha}\pi_{\alpha} + \pi_3\right)\left(\frac{\kappa}{\kappa^*}\right)^2\right], \quad (7)$$

where we have dropped the nucleon rest energy and it is understood that the nucleon coordinate should be substituted in the terms without integrals. From (7), the magnetic moment operator \mathbf{M} is obtained as

$$\mathbf{M} = \frac{e}{2\kappa}\frac{1+\tau_3}{2}\sigma - \frac{e}{2}\int\left[\phi_1\mathbf{x}\times\frac{\partial}{\partial\mathbf{x}}\phi_2 - \phi_2\mathbf{x}\times\frac{\partial}{\partial\mathbf{x}}\phi_1\right]d\mathbf{x} - \frac{ie}{2\kappa}\frac{1+\tau_3}{2}\mathbf{x}\times\frac{\partial}{\partial\mathbf{x}} \\ - \frac{ef}{4\kappa}\left(\frac{\kappa}{\kappa^*}\right)\mathbf{x}\times\sigma(\phi_1\tau_2 - \phi_2\tau_1) + \frac{ef}{16\kappa^2}\mathbf{x}\times\sigma\left[\left(\frac{\kappa}{\kappa^*}\right)^2\left(\sum_{\alpha}\pi_{\alpha}\tau_{\alpha} + \pi_3\right) + \left(\sum_{\alpha}\pi_{\alpha}\tau_{\alpha} + \pi_3\right)\left(\frac{\kappa}{\kappa^*}\right)^2\right]. \quad (8)$$

In this paper we take as our starting point the approximate Hamiltonian (6) and the approximate magnetic moment operator (8).

If we put $\kappa^* = \kappa$ and neglect the nucleon motion in (6), we get the same Hamiltonian as used in the previous investigations¹⁾. It is seen that two essential improvements are done: the first is the nucleon motion, although non-relativistic, and the second is the meson pair term characteristic in *ps-ps* meson theory, though we shall make further approximations in the following relating to these two. It is added that the meson pair effect

will be taken into account so as to give the pair damping automatically, which is intimately connected with the fact that we have not expanded the operator κ^* in (6).

§ 3. Elimination of nucleon coordinate

To calculate the magnetic moments we shall first solve the meson-nucleon problem with the Hamiltonian (6), using the Tomonaga intermediate coupling theory¹⁾ and then make the expectation value of $M(8)$ in terms of the solution obtained, since the electromagnetic interaction may be treated as a weak perturbation. In order to apply the intermediate coupling theory, we first eliminate the nucleon coordinate²⁾, employing the fact that the operator P_{op} , defined by

$$P_{op} = -i \frac{\partial}{\partial \mathbf{x}} - \sum_{\alpha} \int \pi_{\alpha} \frac{\partial \phi_{\alpha}}{\partial \mathbf{x}} d\mathbf{x} \quad (9)$$

commutes with H''' . We can show that the total momentum operator in the original Hamiltonian (1) is transformed into P_{op} after the transformations and approximations we have done to derive the Hamiltonian (5), (6) and (7). Thus we can interpret P_{op} as the total momentum operator of the system. The eigenstate ψ of P_{op} belonging to its eigenvalue \mathbf{P} is given by

$$\psi = (1/\sqrt{V}) \exp \{ i(\mathbf{P} + \sum_{\alpha} \int \pi_{\alpha} \frac{\partial \phi_{\alpha}}{\partial \mathbf{x}} d\mathbf{x}) \cdot \mathbf{x} \} \Psi, \quad (10)$$

where Ψ does not contain the nucleon coordinate \mathbf{x} any longer, V being the normalization volume.

Let us consider the case where the center of mass of the whole system is at rest or $\mathbf{P}=0$. Then the unitary transformation in terms of $\exp [i \{ \sum_{\alpha} \int \pi_{\alpha} (\partial \phi_{\alpha} / \partial \mathbf{x}) d\mathbf{x} \} \cdot \mathbf{x}]$ can eliminate the nucleon coordinate in (6), the result of which is given by H''' in which \mathcal{A} is replaced by $-\left[\sum_{\alpha} \int \pi_{\alpha} (\partial \phi_{\alpha} / \partial \mathbf{x}) d\mathbf{x} \right]^2$ and the nucleon coordinates are put equal to zero in the final two terms without integrals. Then the Schrödinger equation satisfied by Ψ allows a direct application of Tomonaga's method³⁾. It should be noted that only the first two terms of (8) give non-vanishing contributions to the expectation value of \mathbf{M} , since the remaining two vanish under the integration with respect to \mathbf{x} . The magnetic moment can thus be calculated as the expectation value of

$$\mathbf{M}' = \frac{e}{2\kappa} \frac{1+\tau_3}{2} \boldsymbol{\sigma} - \frac{e}{2} \left[\int \phi_1 \mathbf{x} \times \frac{\partial}{\partial \mathbf{x}} \phi_2 - \phi_2 \mathbf{x} \times \frac{\partial}{\partial \mathbf{x}} \phi_1 \right] d\mathbf{x} \quad (11)$$

with respect to Ψ (not Φ). The first and second terms represent the nucleon and meson contributions, respectively. It is seen that the effect of the nucleon recoil does not appear explicitly in the magnetic moment operator (11). Its effect is, however, contained through the terms in the Hamiltonian and can be said as indirect with respect to its influence upon the magnetic moment.

It is valuable here to note that there are two more constants of motion of the

Hamiltonian (6):

$$\mathbf{J} = -i\mathbf{x} \times (\partial/\partial\mathbf{x}) + (1/2)\sigma - \sum_{\alpha} \{\pi_{\alpha}(\mathbf{x} \times \partial/\partial\mathbf{x})\phi_{\alpha} d\mathbf{x}, \quad (12)$$

$$\mathbf{T} = + (1/2)\boldsymbol{\tau} + \int (\boldsymbol{\phi} \times \boldsymbol{\pi}) d\mathbf{x}, \quad (13)$$

in the latter of which we have used vector notations of ϕ_{α} and π_{α} . Quite the same way as \mathbf{P}_{op} in (9), \mathbf{J} and \mathbf{T} can be interpreted as the total angular momentum and isotopic angular momentum operators in units of \hbar . It is clear that \mathbf{P}_{op} commutes with \mathbf{T} but not with \mathbf{J} . Therefore, ψ cannot be an eigenstate of \mathbf{J} , since ψ is an eigenstate of \mathbf{P}_{op} . Let us, however, consider the averaged operator \mathbf{J}' of \mathbf{J} with respect to the nucleon coordinate only:

$$\begin{aligned} \mathbf{J}' &= (1/V) \left\{ \exp \left[i \sum_{\alpha} \{ \pi_{\alpha} (\partial\phi_{\alpha}/\partial\mathbf{x}) d\mathbf{x} \} \mathbf{x} \right] \mathbf{J} \exp \left[-i \sum_{\alpha} \{ \pi_{\alpha} (\partial\phi_{\alpha}/\partial\mathbf{x}) d\mathbf{x} \} \mathbf{x} \right] \right. \\ &\quad \left. + (1/2)\sigma - \sum_{\alpha} \{ \pi_{\alpha}(\mathbf{x} \times \partial/\partial\mathbf{x})\phi_{\alpha} d\mathbf{x} \right\}. \end{aligned} \quad (14)$$

Then \mathbf{T} and \mathbf{J}' both commute with the Hamiltonian acting upon ψ (not ϕ). We shall obtain such a solution of ψ that has eigenvalues $\pm (1/2)$ of operators \mathbf{T} and \mathbf{J}' , which means that our solution corresponds to a nucleon coupled with meson field, which is at rest as a whole ($\mathbf{P}=0$) and has eigenvalues $\pm (1/2)\hbar$ of the total isotopic angular momentum (proton or neutron) and expectation values $\pm (1/2)\hbar$ of the total angular momentum (spin up or down).

§ 4. Spherical wave expansion

For the present purpose, it would be convenient to expand the meson field ϕ_{α} and π_{α} in terms of spherical waves:

$$\begin{aligned} \phi_{\alpha}(\mathbf{x}) &= \sum_{\lambda} (1/2\omega_{\lambda})^{1/2} (a_{\lambda}^{\alpha} + a_{-\lambda}^{\alpha*}) U_{\lambda}(\mathbf{x}), \\ \pi_{\alpha}(\mathbf{x}) &= \sum_{\lambda} i(\omega_{\lambda}/2)^{1/2} (a_{\lambda}^{\alpha*} - a_{-\lambda}^{\alpha}) U_{\lambda}^*(\mathbf{x}), \end{aligned} \quad (15)$$

with

$$\lambda = (k, l, m), \quad -\lambda = (k, l, -m), \quad \omega_{\lambda} = \sqrt{k^2 + \mu^2},$$

where the normalized eigenfunctions $U_{\lambda}(\mathbf{x})$ are given by

$$U_{\lambda}(\mathbf{x}) = (2/R) k j_l(kr) Y_{lm}(\theta, \varphi), \quad (16)$$

normalized within a sphere of the radius R . The allowed eigenvalues of k are $n\pi/R$ or $(n+1/2)\pi/R$ for even or odd l , respectively, where n is zero or positive integer. In (15) and (16), a_{λ}^{α} and $a_{\lambda}^{\alpha*}$ are the annihilation and creation operators of a meson specified by α and λ ; $j_l(kr)$ and $Y_{lm}(\theta, \varphi)$ are the spherical Bessel and spherical harmonic functions.

Then it is shown that, in the Hamiltonian belonging to ψ , only the *S-wave* mesons appear in κ^* and only the *P-wave* ones in the derivative term. In the nucleon recoil term, however, all waves appear and the successive two waves having l -values different by one

are coupled together. It is, therefore, seen that only S - and P -waves interact with a nucleon when the nucleon recoil is neglected; once it is taken into account, all waves appear with mutual coupling terms. We shall, however, retain only S - and P -wave mesons and neglect all others in the nucleon recoil term.

To simplify the notation, we now introduce the S -wave meson annihilation operator $a_{k\alpha}$ and the P -wave one $b_{k\alpha i}$ by

$$\begin{aligned} a_{k\alpha} &\equiv a_{k00}^{\alpha}, & b_{k\alpha 1} &\equiv (1/\sqrt{2}) (a_{k11}^{\alpha} + a_{k1-1}^{\alpha}), \\ b_{k\alpha 2} &\equiv (i/\sqrt{2}) (a_{k11}^{\alpha} - a_{k1-1}^{\alpha}), & b_{k\alpha 3} &\equiv a_{k10}^{\alpha}. \end{aligned} \quad (17)$$

In order to treat the κ^* term in the Hamiltonian, we rewrite it as

$$\kappa^* - \kappa = \sum_{\alpha} f^{\alpha} \phi_{\alpha}^2 / (\kappa + \kappa^*) \quad (18)$$

and then treat the denominator of (18) and also κ^* appearing in the derivative term only approximately, which will be specified in section 6. The Hamiltonian acting on Ψ is then given by, dropping dashes,

$$\begin{aligned} H = & \sum_{k\alpha} \left(\omega_k + \frac{k^2}{2\kappa} \right) a_{k\alpha}^* a_{k\alpha} + \sum_{k\alpha i} \left(\omega_k + \frac{k^2}{2\kappa} \right) b_{k\alpha i}^* b_{k\alpha i} \\ & + [f^2/4\pi R(\kappa + \kappa^*)] \sum_{\alpha} \sum_{kk'} \frac{kk'}{\sqrt{\omega_k \omega_{k'}}} [a_{k\alpha}^* a_{k'\alpha}^* + 2a_{k\alpha}^* a_{k'\alpha} + a_{k\alpha} a_{k'\alpha}] \\ & + [f\kappa/4\sqrt{3\pi R} (\kappa^*)^2] \sum_{k\alpha i} \frac{k^2}{\sqrt{\omega_k}} \sigma_i \tau_{\alpha} [b_{k\alpha i}^* + b_{k\alpha i}], \end{aligned} \quad (19)$$

where we have dropped a constant self energy term and a direct S - P coupling term due to the nucleon recoil*. We should remark, however, that S - and P -terms are never separable in this Hamiltonian because of the κ^* -term. It is added that the P - D cross term in the nucleon recoil can contribute to the nucleon recoil energy in (19).

The same procedure transforms T and J' into the following expressions:

$$T_{\tau} = (1/2) \tau_{\tau} + i \sum_k [(a_{k\alpha} a_{k\beta}^* - a_{k\alpha}^* a_{k\beta}) + \sum_i (b_{i\alpha i} b_{k\beta i}^* - b_{i\alpha i}^* b_{k\beta i})], \quad (20)$$

$$J'_k = (1/2) \sigma_k + i \sum_k \sum_{\alpha} (b_{k\alpha i} b_{k\alpha j}^* - b_{k\alpha i}^* b_{k\alpha j}), \quad (21)$$

where $(\alpha \beta \gamma)$ and $(i j k)$ are even permutations of $(1 2 3)$. As it should be the case, the S -meson does not contribute to J' .

$M'(11)$ then takes a form

$$M'_3 = \frac{e}{2\kappa} \frac{1 + \tau_3}{2} \sigma_3 - \sum_k \frac{e}{2\omega_k} [(b_{k11} + b_{k11}^*) (b_{k22} + b_{k22}^*)]$$

*) The neglect of this S - P coupling term is due to (i) that, according to our more detailed calculation (see the footnote on page 3), it has only a small numerical effect on the final result and (ii) that a great simplification is attained by this neglect when estimating the effect of transitions to excited states; especially equation (34) in section 5 is satisfied only when we neglect this term (see the discussion after (34) in section 5).

$$-(b_{k12} + b_{k12}^*)(b_{k21} + b_{k21}^*)], \quad (22)$$

where the S -meson again does not contribute.

§ 5. Tomonaga's approximation

Following Tomonaga⁵⁾, we expand a and b in terms of the orthogonal functions f and g ;

$$\begin{aligned} a_{k\alpha} &= f_0(k) a_\alpha + \sum_s f_s(k) a_\alpha^s, \\ b_{k\alpha i} &= g_0(k) b_{\alpha i} + \sum_s g_s(k) b_{\alpha i}^s, \end{aligned} \quad (23)$$

where we assume that

$$f_0(k) = \frac{f}{\sqrt{R}} \frac{k}{(\omega_k + k^2/2\kappa) \sqrt{\omega}}, \quad g_0(k) = \frac{g}{\sqrt{R}} \frac{k^2}{(\omega_k + k^2/2\kappa) \sqrt{\omega_k}} \quad (24)$$

with the normalization constants f and g . The functional form of $g_0(k)$ was determined by the general prescription in the intermediate coupling theory and is the same as used by other authors¹⁾ except the appearance of the nucleon recoil energy in the energy denominator of (24). The form of $f_0(k)$ has been determined from an analogous argument, which is explained in Appendix I. It is added that the integral over k should be cut off at some suitable value.

Upon substituting (23) in (19), we get

$$H = H_0 + H_{0s} + H_{gs}, \quad (25)$$

where

$$\begin{aligned} H_0 &= \sum_\alpha [M_\alpha^{00} a_\alpha^* a_\alpha + \sum_i M_b^{0i} b_{\alpha i}^* b_{\alpha i}] + \frac{f^2}{4\pi} \left(\frac{\kappa}{\kappa + \kappa^*} \right) V_\alpha^{00} \sum_\alpha [a_\alpha^* a_\alpha^* \\ &\quad + 2a_\alpha^* a_\alpha + a_\alpha a_\alpha] + \frac{f}{4\sqrt{3}\pi} \left(\frac{\kappa}{\kappa^*} \right)^2 V_b^{0i} \sum_{\alpha i} \tau_\alpha \sigma_i [b_{\alpha i}^* + b_{\alpha i}], \end{aligned} \quad (26)$$

$$\begin{aligned} H_{0s} &= \sum_s \sum_\alpha [M_\alpha^{0s} (a_\alpha^* a_\alpha^s + a_\alpha a_\alpha^{s*}) + \sum_i M_b^{0is} (b_{\alpha i}^* b_{\alpha i}^s + b_{\alpha i} b_{\alpha i}^{s*})] \\ &\quad + \sum_s \frac{f^2}{4\pi} \left(\frac{\kappa}{\kappa + \kappa^*} \right) V_\alpha^{0s} \sum_\alpha [a_\alpha^* a_\alpha^{s*} + a_\alpha^s a_\alpha^* + a_\alpha a_\alpha^{s*} + a_\alpha^s a_\alpha] \\ &\quad + \sum_s \frac{f}{4\sqrt{3}\pi} \left(\frac{\kappa}{\kappa^*} \right)^2 V_b^{0is} \sum_{\alpha i} \tau_\alpha \sigma_i [b_{\alpha i}^{s*} + b_{\alpha i}^s], \end{aligned} \quad (27)$$

$$H_g = \sum_s \sum_\alpha [M_\alpha^{gs} a_\alpha^{s*} a_\alpha^s + \sum_i M_b^{gis} b_{\alpha i}^{s*} b_{\alpha i}^s], \quad (28)$$

with

$$M_\alpha^{ss'} = \sum_k (\omega_k + k^2/2\kappa) f_s(k) f_{s'}(k), \quad M_b^{ss'} = \sum_k (\omega_k + k^2/2\kappa) g_s(k) g_{s'}(k),$$

$$V_{\alpha}^{ss'} = \frac{1}{\kappa} \left(\sum_k \frac{1}{\sqrt{R}} = \frac{k}{\sqrt{\omega_k}} f_s(k) \right) \left(\sum_k \frac{1}{\sqrt{R}} = \frac{k}{\sqrt{\omega_k}} f_{s'}(k) \right),$$

$$V_b^s = \frac{1}{\kappa} \left(\sum_k \frac{1}{\sqrt{R}} = \frac{k^2}{\sqrt{\omega_k}} g_s(k) \right), \quad (29)$$

which are valid for all s including 0. In (28), we have neglected the double unbound meson term due to the third term in (19), since we are going to carry out the lowest order calculation with respect to the unbound mesons. Also we assumed that f and g are real and satisfy

$$\sum_k (\omega_k + k^2/2\kappa) f_s(k) f_{s'}(k) = M_{\alpha}^{ss} \delta_{ss'}, \quad \sum_k (\omega_k + k^2/2\kappa) g_s(k) g_{s'}(k) = M_b^{ss} \delta_{ss'}. \quad (30)$$

We shall furthermore neglect all the unbound meson terms in κ^* . The expansion (23) gives the expressions for $M_{\alpha}'(22)$, $T_{\tau}(20)$ and $J_k'(21)$:

$$M_3 = M_{3,0} + M_{3,0s} + M_{3,ss},$$

$$T_{\tau} = T_{\tau,0} + T_{\tau,s}, \quad J_k' = J_{k,0}' + J_{k,s}', \quad (31)$$

where we do not make any approximations. Naturally no cross terms appear between bound and unbound mesons in T_{τ} and J_k' .

Finally we mention our method to treat unbound mesons. It is most convenient to choose $f_s(k)$ and $g_s(k)$ in (23) as the eigenfunctions of the integral equations⁵⁾

$$(K - S) f_s(k) = f_0(k) \left(\sum_{k'} f_0(k') K' f_s(k') \right),$$

$$(K - S) g_s(k) = g_0(k) \left(\sum_{k'} g_0(k') K' g_s(k') \right), \quad (32)$$

where $K = \omega_k + k^2/2\kappa$ and $S = \omega_s + s^2/2\kappa$. Then the equations (30) are shown to be satisfied, with $M_{\alpha}^{ss} = M_b^{ss} = S$. In practical calculations, we need not specify the particular functional forms of f_s and g_s ; the summation over s can be carried out with the help of the integral equations (32) and the equations such as

$$\sum_s f_s(k) f_s(k') = \delta_{kk'} - f_0(k) f_0(k'). \quad (33)$$

Using the interaction Hamiltonian $H(27)$, we can show that

$$\langle m a_{\alpha}^{\dagger} | H_{0s} | n \rangle = - (M_{\alpha}^{0s} / M_{\alpha}^{00}) (E_m - E_n) \langle m | a_{\alpha} | n \rangle,$$

$$\langle m b_{\alpha}^{\dagger} | H_{0s} | n \rangle = - (M_b^{0s} / M_b^{00}) (E_m - E_n) \langle m | b_{\alpha} | n \rangle, \quad (34)$$

where $|m\rangle$ and $|n\rangle$ are eigenstates of H_0 belonging to eigenvalues E_m and E_n , and $|m a_{\alpha}^{\dagger}\rangle$ is a state in which one unbound meson specified by a_{α}^{\dagger} is present, while the meson configuration is specified by the quantum number m . (34) implies that the bound meson state must always be changed under the virtual emission or absorption of one unbound meson; this simplifies the calculation very much (see the footnote on page 6).

§ 6. Solutions of eigenvalue problem

Now we must specify our approximate method to treat κ^* , which is, after the trans-

formations and approximations thus far introduced, given by

$$\kappa^* = \kappa \left[1 + \sum_{\alpha} (f^2/4\pi\kappa) V_{\alpha}^{(0)} (a_{\alpha}^* a_{\alpha}^* + 2a_{\alpha}^* a_{\alpha} + a_{\alpha} a_{\alpha}) \right]^{1/2}, \quad (35)$$

neglecting unbound mesons and an infinite self energy. In evaluating the matrix element of some operator $f(\kappa^*)$ containing κ^* with respect to states $|i\rangle$ and $|f\rangle$, we approximate $\langle f|f(\kappa^*)|i\rangle$ as $(1/2) \{f(\kappa^*|_f) + f(\kappa^*|_i)\}$, where $\kappa^*|_i$ is $\kappa[1 + \sum_{\alpha} (f^2/4\pi\kappa) V_{\alpha}^{(0)} \langle i|2a_{\alpha}^* a_{\alpha}|i\rangle]^{1/2}$. This procedure is exact when neither $|i\rangle$ nor $|f\rangle$ contains *S-wave* mesons, except the self energy terms. Otherwise, it is but one very simple approximation. If either $|i\rangle$ or $|f\rangle$ contains *S-wave* mesons, the above procedure makes κ^* much larger than κ , which reduces the strength of the interaction since κ^* appears always in the denominator of the interaction Hamiltonian in (26). Thus, our approximate procedure can give rise to the pair damping automatically.

To solve the eigenvalue problem belonging to the Hamiltonian H_0 (26), we expand the state vector in terms of basic vectors of T_{τ} and J_z belonging to eigenvalues $\pm 1/2$ and determine the coefficients so as to minimize the total energy of the system. Retaining amplitudes of up to 3 mesons*), there are 9 independent basic vectors for states (1/2, 1/2) (4 others do not couple to these), 7 for states (3/2, 3/2) and 5 for states (1/2, 3/2) and (3/2, 1/2) where the numbers in brackets are the total isotopic and ordinary angular momentum quantum numbers, respectively. The explicit expressions of basic vectors are given in Appendix II. The secular equations were solved numerically

Table 1. The probability amplitudes, squares of which are normalized to unity, in the ground states (1/2, 1/2), together with the eigenvalues, for various values of the *ps-ps* coupling constant and the cut-off momentum.

| Notation | A_1 | A_2 | A_3 | A_4 |
|--|---------|---------|---------|---------|
| Cut-off(μ) | 6 | 6 | 6 | 5 |
| <i>ps-ps</i> coupling constant ($f/\sqrt{4\pi}$) | 5 | 4 | 3 | 5 |
| Eigenvalue (κ) | -1.5195 | -0.9873 | -0.5566 | -1.0378 |
| $g_1(0, 0)$ | 0.5287 | 0.5866 | 0.6472 | 0.5803 |
| $g_2(0, 1)$ | -0.2916 | -0.2664 | -0.2217 | -0.2581 |
| $g_3(2, 0)$ | -0.1009 | -0.0874 | -0.0686 | -0.1016 |
| $g_4(0, 2)$ | 0.0075 | 0.0063 | 0.0041 | 0.0053 |
| $g_5(0, 2)$ | 0.0264 | 0.0225 | 0.0150 | 0.0190 |
| $g_6(2, 1)$ | 0.0427 | 0.0288 | 0.0155 | 0.0342 |
| $g_7(0, 3)$ | -0.0005 | -0.0004 | -0.0002 | -0.0003 |
| $g_8(0, 3)$ | -0.0014 | -0.0011 | -0.0005 | -0.0008 |
| $g_9(0, 3)$ | 0.0007 | 0.0005 | 0.0003 | 0.0004 |

*) 3 mesons, including both *S*- and *P*-mesons. It is again remarked that H_0 (26) is not separable with respect to *S*- and *P*-mesons because of the κ^* -term.

| A_5 | A_6 | A_7 | A_8 | A_9 |
|---------|---------|---------|---------|---------|
| 5 | 5 | 4 | 4 | 4 |
| 4 | 3 | 5 | 4 | 3 |
| -0.6623 | -0.3624 | -0.6358 | -0.3961 | -0.2088 |
| 0.6435 | 0.7317 | 0.6463 | 0.7117 | 0.7958 |
| -0.2275 | -0.1799 | -0.2141 | -0.1808 | -0.1337 |
| -0.0866 | -0.0652 | -0.0991 | -0.0818 | -0.0576 |
| 0.0041 | 0.0024 | 0.0036 | 0.0023 | 0.0012 |
| 0.0151 | 0.0091 | 0.0120 | 0.0086 | 0.0046 |
| 0.0221 | 0.0111 | 0.0247 | 0.0150 | 0.0069 |
| -0.0002 | -0.0001 | -0.0001 | -0.0001 | -0.0000 |
| -0.0005 | -0.0002 | -0.0004 | -0.0002 | -0.0001 |
| 0.0003 | 0.0001 | 0.0002 | 0.0001 | 0.0000 |

Table 2. The probability amplitudes in the first excited states (1/2, 1/2), squares of which are normalized to unity, together with the eigenvalues, for various values of the ps - ps coupling constant and the cut-off momentum.

| Notation | A_1 | A_2 | A_3 | A_4 |
|--|---------|---------|---------|---------|
| Cut-off (μ) | 6 | 6 | 6 | 5 |
| ps - ps coupling constant ($f/\sqrt{4\pi}$) | 5 | 4 | 3 | 5 |
| Eigenvalue (κ) | -0.1886 | +0.1584 | +0.4289 | -0.0120 |
| a_1 (0, 0) | 0.2564 | 0.2341 | 0.1971 | 0.2305 |
| a_2 (0, 1) | 0.3348 | 0.3745 | 0.4653 | 0.4102 |
| a_3 (2, 0) | -0.0896 | -0.0677 | -0.0426 | 0.0712 |
| a_4 (0, 2) | -0.0541 | -0.0586 | -0.0544 | -0.0449 |
| a_5 (0, 2) | -0.1468 | -0.1626 | -0.1623 | -0.1324 |
| a_6 (2, 1) | -0.0787 | -0.0647 | -0.0506 | -0.0840 |
| a_7 (0, 3) | 0.0066 | 0.0064 | 0.0048 | 0.0046 |
| a_8 (0, 3) | 0.0228 | 0.0218 | 0.0156 | 0.0152 |
| a_9 (0, 3) | -0.0100 | -0.0097 | -0.0072 | -0.0070 |

| A_5 | A_6 | A_7 | A_8 | A_9 |
|---------|---------|---------|---------|---------|
| 5 | 5 | 4 | 4 | 4 |
| 4 | 3 | 5 | 4 | 3 |
| +0.2376 | +0.4305 | +0.1135 | +0.2779 | +0.4025 |
| 0.2038 | 0.1637 | 0.1947 | 0.1648 | 0.1248 |
| 0.4632 | 0.5633 | 0.5041 | 0.5682 | 0.6708 |
| -0.0513 | -0.0295 | -0.0505 | -0.0337 | -0.0171 |

| | | | | |
|---------|---------|---------|---------|---------|
| -0.0468 | -0.0415 | -0.0355 | -0.0352 | -0.0293 |
| -0.1419 | -0.1341 | -0.1135 | -0.1158 | -0.1017 |
| -0.0691 | -0.0518 | -0.0856 | -0.0690 | -0.0485 |
| 0.0042 | 0.0029 | 0.0029 | 0.0024 | 0.0014 |
| 0.0135 | 0.0089 | 0.0089 | 0.0073 | 0.0043 |
| -0.0064 | -0.0043 | -0.0043 | -0.0036 | -0.0022 |

Table 3. The probability amplitudes in the ground state $(3/2, 3/2)$, squares of which are normalized to unity, together with the eigenvalues, for various values of the ps - ps coupling constant and the cut-off momentum.

| Notation | A_1 | A_2 | A_3 | A_4 |
|---|---------|---------|---------|---------|
| Cut-off (μ) | 6 | 6 | 6 | 5 |
| ps - ps coupling constant ($f/\sqrt{4\pi}$) | 5 | 4 | 3 | 5 |
| Eigenvalue (κ) | -0.7178 | -0.2469 | +0.1470 | -0.3689 |
| b_1 (0, 1) | 0.6572 | 0.6744 | 0.7144 | 0.6892 |
| b_2 (0, 2) | -0.0165 | -0.0172 | -0.0153 | -0.0134 |
| b_3 (2, 0) | -0.1811 | -0.1925 | -0.1883 | -0.1637 |
| b_4 (2, 1) | -0.1264 | -0.0977 | -0.0681 | -0.1200 |
| b_5 (0, 3) | 0.0046 | 0.0045 | 0.0035 | 0.0034 |
| b_6 (0, 3) | 0.0004 | 0.0004 | 0.0003 | 0.0003 |
| b_7 (0, 3) | 0.0138 | 0.0133 | 0.0101 | 0.0099 |

| A_5 | A_6 | A_7 | A_8 | A_9 |
|---------|---------|---------|---------|---------|
| 5 | 5 | 4 | 4 | 4 |
| 4 | 3 | 5 | 4 | 3 |
| -0.0321 | +0.2462 | -0.1000 | +0.1189 | +0.2965 |
| 0.6892 | 0.7566 | 0.7288 | 0.7567 | 0.8055 |
| -0.0134 | -0.0114 | -0.0103 | -0.0098 | -0.0078 |
| -0.1700 | -0.1602 | -0.1427 | -0.1272 | -0.1272 |
| -0.0924 | -0.0631 | -0.1098 | -0.0834 | -0.0547 |
| 0.0031 | 0.0023 | 0.0023 | 0.0019 | 0.0013 |
| 0.0003 | 0.0002 | 0.0002 | 0.0002 | 0.0001 |
| 0.0090 | 0.0063 | 0.0064 | 0.0054 | 0.0034 |

Table 4. The probability amplitudes in the ground states $(1/2, 3/2)$ or $(3/2, 1/2)$, squares of which are normalized to unity, together with the eigenvalues, for various values of the ps - ps coupling constant and the cut-off momentum.

| Notation | A_1 | A_2 | A_3 | A_4 |
|--|---------|---------|---------|---------|
| Cut-off (μ) | 6 | 6 | 6 | 5 |
| ps - ps coupling constant ($f/\sqrt{4\pi}$) | 5 | 4 | 3 | 5 |
| Eigenvalue (κ) | -0.3265 | +0.0840 | +0.3962 | -0.1072 |
| c_1 (0, 1) | 0.7766 | 0.8078 | 0.8519 | 0.7998 |
| c_2 (0, 2) | -0.0435 | -0.0488 | -0.0481 | -0.0377 |
| c_3 (2, 1) | -0.1788 | -0.1425 | -0.0993 | -0.1620 |
| c_4 (0, 3) | 0.0002 | 0.0002 | 0.0001 | 0.0001 |
| c_5 (0, 3) | 0.0007 | 0.0007 | 0.0005 | 0.0004 |

| A_5 | A_6 | A_7 | A_8 | A_9 |
|---------|---------|---------|---------|---------|
| 5 | 5 | 4 | 4 | 4 |
| 4 | 3 | 5 | 4 | 3 |
| +0.1857 | +0.4066 | +0.0571 | +0.2468 | +0.3879 |
| 0.8317 | 0.8480 | 0.8277 | 0.8604 | 0.9033 |
| -0.0406 | -0.0378 | -0.0312 | -0.0317 | -0.0275 |
| -0.1270 | -0.0858 | -0.1408 | -0.1076 | -0.0691 |
| 0.0001 | 0.0001 | 0.0001 | 0.0001 | 0.0000 |
| 0.0004 | 0.0003 | 0.0003 | 0.0002 | 0.0001 |

by the digital electronic computer BESK (Binary Electronic Sequence Computer) in Stockholm, Sweden. The results of the calculations are summarized in Tables 1, 2, 3 and 4. The former and the latter figures in brackets in the Tables represent the numbers of S - and P -wave mesons, respectively, in respective basic vectors.

Some characteristic points of these results are summarized: First, the approximation of retaining up to 3 mesons may be regarded as reasonable, especially for the ground state $(1/2, 1/2)$. Secondly, the S -meson amplitudes are seen to be reduced rather much and be comparable with the corresponding P -meson amplitudes. Thirdly, the eigenvalue depends rather sensitively on $f/\sqrt{4\pi}$ and cut-off; the probabilities of various states are, however, rather insensitive to them, which implies that the calculated magnetic moment would also be insensitive to them. Finally, the isobar separation seems to be somewhat too large to fit the pion-nucleon scattering resonance.

§ 7. Magnetic moments

a) Magnetic moments without unbound meson corrections

The magnetic moment can be calculated straightforwardly, by evaluating the expectation value of M_z given by (22) and (31). Let us first neglect all the unbound meson effects; the expectation value of $M_{z,0}$ in terms of the ground state solution is given, in units of $e/2\kappa$, by

$$\mu_P = 1 - \frac{10}{9}A + B, \quad \mu_N = -\frac{2}{9}A - B, \quad (36)$$

where A and B are functions of the probability amplitudes of various states:

$$\begin{aligned} A &= g_2^2 + g_5^2 + g_6^2 + g_8^2 + g_9^2 = 1 - (g_1^2 + g_3^2 + g_4^2 + g_7^2), \\ B &= \left(\frac{\kappa}{\mu}\right) \left(\sum_k \mu \frac{g^2(k)}{\omega_k}\right) \left(\frac{4}{\sqrt{18}}g_1g_5 + \frac{4}{9}g_2^2 + \frac{8}{9}g_4g_5 + \frac{2}{9}g_5^2 + \frac{4}{3\sqrt{6}}g_2g_7 + \frac{4}{9}g_6^2 \right. \\ &\quad \left. + \frac{8}{3\sqrt{33}}g_7g_8 - \frac{20}{3\sqrt{66}}g_7g_9 + \frac{16}{9\sqrt{22}}g_2g_3 - \frac{20}{9\sqrt{11}}g_2g_9 + \frac{20}{33}g_3^2 - \frac{40\sqrt{2}}{99}g_8g_9 - \frac{3}{11}g_9^2\right), \end{aligned} \quad (37)$$

where g_i 's are the probability amplitudes in the ground state, given in Table 1. From (36), it is readily seen that

$$\mu_P + \mu_N = 1 - \frac{4}{3}A, \quad (38)$$

which is just Sachs' theorem⁽⁶⁾, since, from (A.6) in the Appendix II, A is the probability that pions (any number or charge) occur in the ground state with the total angular momentum one. If this is equated to the empirical value, 0.8797, then A must be 0.0902, which implies a weak coupling of mesons with nucleons. Our present calculation never predicts so small A (in case A_1 , A is 0.37 and even in case A_0 , it is 0.15). It is expected that, if the theory could predict so small A , then B would correspondingly be small, which would result in too small μ_P and μ_N . Since, however, B contains a divergent integral, it may be possible to make B large enough to explain μ_P and μ_N , without making A increase so much, simply by letting the cut-off momentum large enough. We must stress that such a procedure has no meaning since we based ourselves on the non-relativistic Hamiltonian. Besides this basic difficulty, our numerical results in Table 5 give too small meson contributions. It is finally added that the results are rather insensitive to both coupling constant and cut-off momentum.

Table 5. The calculated values of magnetic moments, for three cases classified in Table 1.

| Experiment | $\mu_P = 2.79$ | | | $\mu_N = -1.91$ | | |
|-----------------------------------|----------------|-------|-------|-----------------|-------|-------|
| Case | A_1 | A_5 | A_9 | A_1 | A_5 | A_9 |
| Nucleon Contribution | 0.58 | 0.70 | 0.84 | -0.08 | -0.06 | -0.03 |
| Meson Contribution | 0.61 | 0.48 | 0.31 | -0.61 | -0.48 | -0.31 |
| Total without Isobar Contribution | 1.19 | 1.18 | 1.15 | -0.69 | -0.54 | -0.34 |
| Total with Isobar Contribution | 1.18 | 1.18 | 1.15 | -0.69 | -0.53 | -0.34 |

b) Effects of excited states

The difficulty relative to Sachs' theorem (38) is nothing but the one which was inferred from a simple intuitive argument by one of the authors⁷⁾. In that paper, he pointed out that the idea of the nucleon isobar may resolve this difficulty, although it was shown afterwards that it might not be the case⁸⁾. Anyway, as the Tomonaga approximation predicts nucleon excited states such as given in Tables 2, 3 and 4, it seems interesting to see how these excited states may modify the results. We have, therefore, extended our calculation so as to take account of the virtual excitation of nucleons under a virtual emission of one unbound meson, where we have considered all the isobar states given in Tables 2, 3, and 4.

For this purpose we have solved the Tamm-Dancoff equation belonging to the total Hamiltonian H given by (25), (26), (27) and (28), retaining up to the one unbound meson amplitude and calculated the expectation value of M_z (31) without any further approximations. Details are given in Appendix III. The results are also given in Table 5. It is seen that the results are still worse, though the correction is quite negligible. The reduction of the calculated values is due to the fact that the isobar effect is insufficient to recover the reduction of the original value due to the change of the normalization.

§ 8. Conclusions

Although not yet quite satisfactorily, we have taken account of the effect of nucleon recoil and *S-wave* mesons partially, so as to give the pair damping automatically, in applying the Tomonaga intermediate coupling meson theory to the symmetrical *ps-ps* meson theory. We have considered also some of the excited states, into which the ground state nucleon can jump under the virtual emission of one unbound meson. The calculated values of nucleon magnetic moments turned out, however, to be too small to be consistent with empirical values for the reasonable coupling constant and cut-off momentum, mainly due to meson contributions. Sachs' theorem has been obtained as regards the bound meson contribution. The effect of excited states has been shown to be quite negligible. The results of these calculations seem to imply that the present meson field theory is essentially insufficient to predict the correct values of nucleon magnetic moments.

Finally one of the authors (M.S.) would like to express his gratitude to Professor L. Hulthén for the hospitality and helpful arrangements in carrying out most of the numerical works in this paper.

Appendix I

In order to find out a reasonable functional form of $f_0(k)$, we shall employ the fact that, if we regard κ^* as a pure constant, the Hamiltonian H_0 given by (26) becomes separable as regards *S-* and *P-wave* mesons, the *S-wave* part of which can be solved rigorously, since it reduces to a system of harmonic oscillators if we use a new set of canonical variables

$$\alpha_a = (a_a + a_a^*) / \sqrt{2}, \quad p_a = (a_a - a_a^*) / \sqrt{2} i. \quad (A.1)$$

The lowest eigenvalue is given by

$$E_0 = (3/2) M_a^{00} \left[\{1 + (2\alpha^2 V_a^{00} / M_a^{00})\}^{1/2} - \{1 + (\alpha^2 V_a^{00} / M_a^{00})\} \right]. \quad (A.2)$$

We then require that E_0 should be minimum for a reasonable choice of $f_0(k)$ or

$$\delta_f E_0 = 0. \quad (A.3)$$

If we solve this variational equation under the restriction that

$$\sum_k f_0^2(k) = 1, \quad (A.4)$$

we get

$$f_0(k) \propto k / [\sqrt{\omega_k} \{ \omega_k + (k^2/2\mu) + \text{const.} \}]. \quad (A.5)$$

In the text we have simply neglected the constant term in the denominator only because it makes the calculation much simpler.

Appendix II

The basic vectors are constructed by using the well-known Clebsch-Gordan's rule.

For $J=1/2$, $J_z=1/2$, $T=1/2$, $T_3=1/2$;

$$\begin{aligned} \Psi_{00}^{00}(1) &= |+, \uparrow\rangle, & \Psi_{11}^{10}(2) &= \sum_{m+\sigma/2=1/2} \sum_{\zeta+\tau/2=1/2} C_{\zeta\tau}^{1/2} C_{m\sigma}^{1/2} b_{\zeta m}^* |\tau, \sigma\rangle, \\ \Psi_{00}^{02}(3) &= A_0^{(2)*} |+, \uparrow\rangle, & \Psi_{00}^{20}(4) &= (1/\sqrt{18}) B_0^{(2)*} |+, \uparrow\rangle, \\ \Psi_{11}^{20}(5) &= \sum_{m+\sigma/2=1/2} \sum_{\zeta+\tau/2=1/2} C_{\zeta\tau}^{1/2} C_{m\sigma}^{1/2} B_{1,\zeta m}^{(2)*} |\tau, \sigma\rangle, \\ \Psi_{11}^{12}(6) &= A_0^{(2)*} \Psi_{11}^{10}(2), & \Psi_{00}^{30}(7) &= (1/\sqrt{6}) \det |b_{\zeta m}^*| |+, \uparrow\rangle, \\ \Psi_{11}^{30}(8) &= (1/\sqrt{22}) B_0^{(2)*} \Psi_{11}^{10}(2), \\ \Psi_{11}^{30}(9) &= (2\sqrt{2}/5) \Psi_{11}^{30}(8) + (\sqrt{11}/5) \sum_{m+\sigma/2=1/2} \sum_{\zeta+\tau/2=1/2} C_{m\sigma}^{1/2} C_{\zeta\tau}^{1/2} B_{1,\zeta m}^{(3)*} |\tau, \sigma\rangle, \end{aligned} \quad (A.6)$$

where

$$\begin{aligned} A_0^{(2)*} &= (1/\sqrt{6}) \sum_{\zeta} a_{\zeta}^* a_{-\zeta}^*, & B_0^{(2)*} &= \sum_{\zeta, m} b_{\zeta m}^* b_{-\zeta-m}^*, \\ B_{1,\zeta m}^{(2)*} &= (1/\sqrt{2}) \sum_{\zeta' + \zeta'' = \zeta} \sum_{m' + m'' = m} d_{\zeta'\zeta''}^1 d_{m'm''}^1 b_{\zeta'm'}^* b_{\zeta''m''}^*, \\ B_{1,\zeta m}^{(3)*} &= \sum_{\zeta' + \zeta'' = \zeta} \sum_{m' + m'' = m} d_{\zeta'\zeta''}^1 d_{m'm''}^1 b_{\zeta'm'}^* B_{1,\zeta'm''}^{(2)*}, \\ C_{\zeta\sigma}^{1/2} &= \langle j_1=1, m_1=\zeta, j_2=1/2, m_2=\sigma/2 | j_1, j_2, j=1/2, m=1/2 \rangle, \\ d_{\zeta'\zeta''}^1 &= \langle j_1=1, m_1=\zeta, j_2=1, m_2=\zeta' | j_1, j_2, j=1, m=\zeta+\zeta' \rangle. \end{aligned}$$

The upper suffixes of Ψ mean the P -wave and S -wave meson numbers and the lower ones denote eigenvalues of the total angular momentum and the total isotopic spin of the meson field.

For $J=3/2$, $J_z=3/2$, $T=1/2$, $T_3=1/2$;

$$\begin{aligned}\Psi_{11}^{10}(1) &= \sum_{\zeta+\tau/2=3/2} C_{\zeta\tau}^{3/2} b_{\zeta 1}^* |\tau, \uparrow\rangle, \quad \Psi_{11}^{20}(2) = \sum_{\zeta+\tau/2=3/2} C_{\zeta\tau}^{3/2} B_{\zeta 1}^{(2)*} |\tau, \uparrow\rangle, \\ \Psi_{11}^{12}(3) &= A_0^{(2)*} \Psi_{11}^0(1), \quad \Psi_{11}^{30}(4) = (1/\sqrt{22}) B_0^{(2)*} \Psi_{11}^{10}(1), \\ \Psi_{11}^{30}(5) &= (2\sqrt{2}/5) \Psi_{11}^{30}(4) + (\sqrt{11}/5) \sum_{\zeta+\tau/2=3/2} C_{\zeta\tau}^{3/2} B_{\zeta 1}^{(3)*} |\tau, \uparrow\rangle,\end{aligned}\quad (A.7)$$

where

$$C_{\zeta\tau}^{3/2} = \langle j_1=1, m_1=\zeta, j_2=1/2, m_2=\tau | j_1, j_2, j=3/2, m=\zeta+\tau/2 \rangle.$$

For $J=1/2$, $J_z=1/2$, $T=3/2$, $T_3=3/2$, we have only to exchange the angular momentum variables and the isotopic spin variables.

For $J=3/2$, $J_z=3/2$, $T=3/2$, $T_3=3/2$;

$$\begin{aligned}\Psi_{11}^{10}(1) &= b_{11}^* |+, \uparrow\rangle, \quad \Psi_{11}^{20}(2) = B_{11}^{(2)*} |+, \uparrow\rangle, \\ \Psi_{22}^{20}(3) &= \sum_{\lambda+\tau/2=3/2} \sum_{\mu+\sigma/2=3/2} f_{\lambda\tau}^{3/2} f_{\mu\sigma}^{3/2} B_{2,\lambda\mu}^{(2)*} |\tau, \sigma\rangle, \\ \Psi_{11}^{12}(4) &= b_{11}^* A_0^{(2)*} |+, \uparrow\rangle, \quad \Psi_{11}^{30}(5) = (1/\sqrt{22}) b^* B^{(3)*} |+, \uparrow\rangle, \\ \Psi_{11}^{30}(6) &= (2\sqrt{2}/5) \Psi_{11}^{30}(5) + (\sqrt{11}/5) B_{11}^{(3)*} |+, \uparrow\rangle, \\ \Psi_{22}^{30}(7) &= \sum_{\lambda+\tau/2=3/2} \sum_{\mu+\sigma/2=3/2} f_{\lambda\tau}^{3/2} f_{\mu\sigma}^{3/2} B_{2,\lambda\mu}^{(3)*} |\tau, \sigma\rangle,\end{aligned}\quad (A.8)$$

where

$$\begin{aligned}B_{2,\lambda\mu}^{(2)*} &= n(\lambda, \mu) \sum_{\zeta+\zeta'=\lambda} \sum_{m+m'=\mu} d_{\zeta\zeta'}^2 d_{mm'}^2 b_{\zeta m}^* b_{\zeta' m'}^*, \\ B_{2,\lambda\mu}^{(3)*} &= n'(\lambda, \mu) \sum_{\zeta+\zeta'=\lambda} \sum_{m+m'=\mu} d_{\zeta\zeta'}^2 d_{mm'}^2 B_{\zeta m}^{(2)*} b_{\zeta' m'}^*, \\ d_{\zeta\zeta'}^2 &= \langle j_1=1, m_1=\zeta, j_2=1, m_2=\zeta' | j_1, j_2, j=2, m=\zeta+\zeta' \rangle, \\ f_{\lambda\tau}^{3/2} &= \langle j_1=2, m_1=\lambda, j_2=1/2, m_2=\tau/2 | j_1, j_2, j=3/2, m=\lambda+\tau/2 \rangle.\end{aligned}$$

$n(\lambda, \mu)$ and $n'(\lambda, \mu)$ are the normalization factors, which are necessary because of dealing with the assembly of identical particles.

Appendix III

To estimate the effects of excited states we solved the Tamm-Dancoff equation

$$\begin{aligned}E|\Psi\rangle &= (H_0 + H_s + H_c) |\Psi\rangle, \\ |\Psi\rangle &= C_0 |g\rangle + \sum_{i,s} C(i, s) |i, s\rangle,\end{aligned}\quad (A.9)$$

retaining only one unbound meson amplitude, where $|g\rangle$ is the ground state of the nucleon and $|i, s\rangle$'s represent the excited states with one unbound meson. The parameter i specifies the state; $i=1, 2, 3$ or 4 corresponds to $J=1/2$, $T=1/2$; $J=3/2$, $T=1/2$; $J=1/2$, $T=3/2$ or $J=3/2$, $T=3/2$, respectively. We do not need the amplitude of the ground state with one unbound meson because of what was discussed in the end of the section 5.

After some elementary calculations, the coefficient $C(i, s)$ and the expectation value of the magnetic moment are given by;

$$C(i, s) = (\partial_i X(i) M_b^{is} / M_b^{i0} (S + \delta_i)) C_{0s}$$

$$\langle \Psi | M_z | \Psi \rangle = \langle \Psi | M_{z,0} | \Psi \rangle + \langle \Psi | M_{z,0s} | \Psi \rangle + \langle \Psi | M_{z,s} | \Psi \rangle, \quad (A \cdot 10)$$

$$\langle \Psi | M_{z,0} | \Psi \rangle = C_0^2 \langle g | M_{z,0} | g \rangle + C_0^2 \sum_{ij} \frac{\partial_i \partial_j X(i) X(j)}{(M_b^{i0})^2} \left[\sum_s \frac{(M_b^{is})^2}{(S + \delta_i)(S + \delta_j)} \right] \times \langle i, s | M_{z,0} | j, s \rangle,$$

$$\langle \Psi | M_{z,0s} | \Psi \rangle = 2 C_0^2 \sum_i \frac{\partial_i X(i)}{M_b^{i0}} \left(\frac{\kappa}{\mu} \right) \left[\sum_s \frac{M_b^{is}}{(S + \delta_i)} \left\{ \sum_k \frac{\mu g_0(k) g_s(k)}{\omega_k} \right\} \right] \times \langle i, s | M_{z,0s} | g \rangle,$$

$$\langle \Psi | M_{z,s} | \Psi \rangle = \frac{4}{9} P(1) - \frac{2}{9} P(2) - \frac{2}{9} P(3) + \frac{1}{9} P(4),$$

where

$$X(i) = \sum_{m+J_z=\lambda} \sum_{\zeta+T_3=\mu} C_{mJ_z}^{J,1/2} C_{\zeta T_3}^{T,1/2} \langle i; J, J_z; T, T_3 | b_{\zeta m} | g; 1/2, \lambda; 1/2, \mu \rangle,$$

$$P(i) = C_0^2 \left(\frac{\kappa}{\mu} \right) \frac{\partial_i X(i)^2}{(M_b^{i0})^2} \left[\sum_{ss'} \frac{M_b^{is} M_b^{is'}}{(S + \delta_i)(S' + \delta_i)} \left(\sum_k \frac{\mu g_s(k) g_{s'}(k)}{\omega_k} \right) \right].$$

$C_{mJ_z}^{J,1/2}$ and $C_{\zeta T_3}^{T,1/2}$ are the Clebsch-Gordan's coefficients and δ_i is the energy difference between the respective excited and ground states. The summation over s can be carried out by virtue of (34).

References

- 1) F. H. Harlow and B. A. Jacobsohn, Phys. Rev. **93**(1954), 333.
G. Takeda, Phys., Rev. **95**(1954), 1078. E. Yamada and S. Hirokawa, Soryusiron Kenkyu (1954), 479 (Mimeographed circular in Japanese).
- 2) L. L. Foldy, Phys. Rev. **84**(1951), 168. J. M. Berger, L. L. Foldy and R. K. Osborn, Phys. Rev. **87**(1952), 1061.
- 3) L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**(1950), 29. R. K. Osborn, Phys. Rev. **86**(1952), 340.
- 4) T. D. Lee and D. Pines, Phys. Rev. **88**(1952), 960. T. D. Lee, F. E. Low and D. Pines, Phys. Rev. **90**(1953), 297. T. D. Lee and D. Pines, Phys. Rev. **92**(1953), 883.
- 5) S. Tomonaga, Prog. Theor. Phys. **2**(1947), 6.
- 6) R. G. Sachs, Phys. Rev. **87**(1952), 1100. W. G. Hollady, Phys. Rev. **101**(1956), 1198. H. Miyazawa, Phys. Rev. **97**(1955), 1413; **101**(1956), 1964.
- 7) M. Sugawara, Prog. Theor. Phys. **8**(1947), 549.
- 8) F. J. Belinfante, Phys. Rev. **92**(1953), 994, 997.

On Quantum-Mechanical Nuclear Dipole Vibrations

Jun-ichi FUJITA

Department of Physics, University of Tokyo, Tokyo

(Received February 28, 1956)

The semi-classical collective model of nuclear photo-effect, which assumes the existence of nuclear dipole vibration, is reformulated quantum-mechanically according to the S. Tomonaga's general prescription, and a modified collective model is proposed, which takes into account the velocity-dependent character of neutron-proton exchange forces. The experimental and theoretical evidences for the validity of the dipole vibration model are discussed.

§ 1. Introduction

Since the giant resonance in nuclear photoeffect was first interpreted by M. Goldhaber and E. Teller¹⁾ as the dipole vibration of the nucleus, in which the bulk of protons and neutrons move in opposite directions, the phenomenon has been studied theoretically by many authors.^{2),3),4)} Various explanations proposed so far can be grossly classified into the following three categories; the independent particle model, the sub-unit model and the collective model. Especially the recent shell model treatments⁵⁾ based on the success of cloudy crystal ball model in the high energy region seem to be successful in interpreting at least qualitatively many details of the experimental facts concerning the giant resonance, i.e., the angular distribution of the emitted particles, the anomalously large yield of protons, the difference between the maxima of the (γ, n) and the (γ, p) excitation curve in some light nuclei, the systematic behaviour of the widths of the giant resonance, etc. However, the difficulties in providing the explanation for the absolute magnitude and the A -dependence of the photon energy corresponding to the maximum absorption cross section and the harmonic mean energy of the excitation curve suggests the necessity of studying some other models involving correlations among the nucleons.

In view of this unsatisfactory situation it would be desirable to investigate the other extreme model, that is, the collective model of photo-reaction, which assumes the existence of nuclear dipole vibration. In most of the previous collective models the dipole vibration was treated semi-classically and it is not clear how to get the knowledge about the detailed structure of giant resonance. We hope that the experimental facts for which the shell model treatment is successful can also be explained by the dipole vibration model if we derive this model quantum-mechanically. Already M. Ferentz, M. Gell-Mann and D. Pines⁶⁾ treated the dipole vibration quantum-mechanically applying D. Bohm and D. Pines' theory of plasma oscillations, but it seems difficult to extend the theory to the problem with the more realistic boundary condition involving the nuclear surface. I. Bloch and

Y. Hsieh²⁾ have also discussed the normal coordinate corresponding to the dipole vibration in a very special case where internucleon forces obey Hooke's law.

Recently quite an interesting collective model has been proposed by S. Takagi and S. Fujii³⁾ and S. Gallone et al.,³⁾ which unifies the existence of the dipole vibration and the validity of the shell model. However, since it is not formulated in a completely quantum-mechanical manner, we do not know how to improve its approximation further and what kind of role the exchange forces in the nucleus play, which may not be negligible.

Recently S. Tomonaga⁶⁾ has given a method to formulate the collective motions quantum-mechanically, which are described by classical displacement potentials. According to his general prescription we can also reformulate the classical dipole vibration quantum-mechanically. The aims of this paper are first to reformulate the dipole vibration model according to his general prescription and secondly to make the physical meaning of the neutron-proton exchange effects clear and construct a collective model which takes into account their effects.

§ 2. Collective coordinate and collective momentum

The quantum-mechanics of a particle interacting with electromagnetic field is well-established and can be applied directly to the theory of giant resonance, in which the electric dipole interaction is known to be predominant,⁷⁾ as is also inferred from the order of the photon wave length $\lambda \cong 6 \times 10^{-12}$ cm or $\hbar\lambda \cong 1 \times 10^{-12}$ cm corresponding to 20 Mev. Thus the situation is quite similar to the problem of light absorption by an atom. There are, however, two points which need special attention. 1) Instead of the true charge e of the proton, we must take the effective charges eN/A and $-eZ/A$ for the proton and the neutron respectively because of their motion relative to the centre of mass of the nucleus. 2) The exchange current between the nucleons, which has no classical analogy, has to be included in the interaction. Of these the second gives rise to a difficulty. A. J. F. Siegert⁸⁾ showed that the usual interaction expressed in terms of ω , the frequency, or the energy difference between the initial and the final states, remains valid in long wave length limit, even though the relation $p = m\dot{z} = im\omega z$, z : the displacement along the electric vector, is no longer tenable. According to L. L. Foldy,¹⁰⁾ Siegert's results may not be utterly correct. For the want of more reliable theory and the consistency¹¹⁾ with the experimental facts about the deuteron disintegration in the energy region of the order of 20 Mev., we take in phenomenological or semi-empirical sense,

$$H_{\text{rad.int.}} = (ie/c) \omega \sum_i \{ (N/A) ((1 - \tau_{zi})/2) - (Z/A) ((1 + \tau_{zi})/2) \} z_i \quad (2 \cdot 1)$$

as the interaction energy between the nucleons and the electric field polarized in z -direction.

Now according to S. Tomonaga's general prescription⁶⁾ we can determine the collective coordinates ξ 's, if the displacement potential which describes the classical collective motion is given. Though the choice of the displacement potential is not unique, it is necessary

that $H_{\text{rad.int.}}$ is expressed either only or partly in the coordinate ξ , which is derived from the displacement potential. For instance we take M. Goldhaber and E. Teller's third model,¹⁾ where the displacement potentials, ϕ_p and ϕ_N for protons and neutrons are

$$\phi_p(\mathbf{r}) = (1/Z)z, \quad \phi_N(\mathbf{r}) = -(1/N)z \quad (2.2)$$

respectively, if we take the stream lines parallel to z -axis. (Fig. 1). We can replace (2.2) by

$$\phi(\mathbf{r}) = \{ (1/Z)(1 - \tau_z)/2 - (1/N)(1 + \tau_z)/2 \} z \equiv tz. \quad (2.3)$$

Then the corresponding collective coordinate ξ is, by the general prescription,

$$\xi = \sum_i \phi(\mathbf{r}_i) = \sum_i t_i z_i. \quad (2.4)$$

Obviously the physical meaning of ξ is the z -component of the distance vector between the centers of mass of neutrons and protons.

Thus (2.1) becomes

$$H_{\text{rad.int.}} = (iew/c)(NZ/A)\xi. \quad (2.5)$$

In this case the electric dipole interaction is expressed in terms of the collective coordinate ξ only. We will restrict our discussions only to the Goldhaber-Teller model in this paper, since it is the simplest type of dipole vibration. The other collective models will be touched upon in Section 10.

The momentum, π , canonically conjugate to ξ , is

$$\pi = -i\hbar \sum_i (NZ/A) t_i \partial / \partial z_i. \quad (2.6)$$

Between ξ and π the commutation relation holds exactly:

$$[\pi, \xi] = -i\hbar. \quad (2.7)$$

Obviously π is the relative momentum of neutrons and protons in the dipole vibration. Since our space is three-dimensional, we need three ξ 's and three π 's to describe the collective motion, though they are not explicitly introduced for simplicity.

Using the canonical pair π and ξ , the collective part corresponding to the dipole vibration is separated from the total Hamiltonian of nucleus according to S. Tomonaga's general procedure as shown in the next section.

§ 3. Separation of the dipole vibration

The total Hamiltonian is

$$H = H_N + H_{\text{rad.}} + H_{\text{rad.int.}}, \quad (3.1)$$

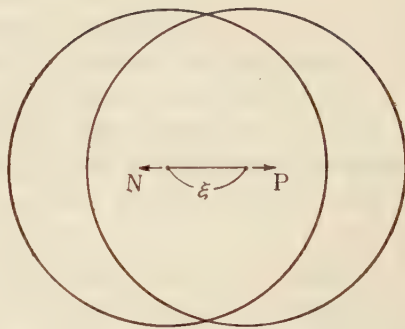


Fig. 1. Goldhaber-Teller model.

where H_N is the Hamiltonian of the nucleus and $H_{\text{rad.}}$ of the radiation field:

$$H_N = -1/2m \cdot \sum_i J_i + V_N. \quad (3.2)$$

For simplicity we assume that V_N consists of the two-body interactions only, the central forces of the Wigner type and the Majorana type. It is quite straightforward to extend the formalism to include the other types of interactions. We also neglect the contribution from tensor forces which may probably be small.

The kinetic energy part of the Hamiltonian can be separated⁽⁵⁾ into the collective part $\pi^2/2M$ and the internal part* containing neither ξ nor π . M represents the reduced mass related to the dipole vibration, that is, $M = NZ/Am$ if we denote the single nucleon mass as m .

The potential part of the Hamiltonian commuting with ξ can be expanded into the power series of ξ in accordance with the general prescription,⁽⁶⁾ while the other non-commuting part, V' , which contains τ_x or τ_y in our case, must be omitted in the course of expansion. The details about V' will be discussed in Section 8. Thus we separate the potential part, V_N into the two parts:

$$V_N \equiv V + V' = \sum V(r_{ij}) + \sum V'(r_{ij}), \quad (3.3)$$

where

$$V(r_{ij}) = V_0(r_{ij}) [1 - x - x \{ (1 + \sigma_i \sigma_j) / 2 \} (1 + \tau_{xi} \tau_{xj}) / 2] \quad (3.3a)$$

and

$$V'(r_{ij}) = -x V_0(r_{ij}) \{ (\tau_{xi} \tau_{xj} + \tau_{yi} \tau_{yj}) / 2 \} (1 + \sigma_i \sigma_j) / 2. \quad (3.3b)$$

The total Hamiltonian of the nucleus can be written as

$$H_N = H^{(0)}(\zeta) + H^{(1)}(\zeta) \xi + (1/2) H^{(2)}(\zeta) \xi^2 + \dots + \pi^2/2M + V'. \quad (3.4)$$

In this expansion, the coefficients $H^{(i)}(\zeta)$ etc., which contain only the internal coordinates ζ 's orthogonal to the collective coordinates ξ 's, can be easily calculated by the general prescription.⁽⁶⁾ In our case

$$\begin{aligned} H^{(1)}(\zeta) &= T^{(1)} + (1-x)V^{(1)} = (1-x)V^{(1)} \\ H^{(2)}(\zeta) &= T^{(2)} + (1-x)V^{(2)} = (1-x)V^{(2)} \end{aligned} \quad (3.5)$$

where $T^{(i)}$ is the coefficient for the i -th power of ξ in the expansion of the internal kinetic energy part, $-(1/2m) \sum_i J_i - \pi^2/2M$, while $(1-x)V^{(i)}$ is that of the potential energy part V . Their explicit expressions are

$$\begin{aligned} (1-x)V^{(1)} &= i[\pi/\hbar, V] - \dot{\xi}^2 [\pi/\hbar, [\pi/\hbar, V]] + \dots \\ &= (NZ/A) \sum \partial V / \partial r_{ij} \{ (z_i - z_j) / r_{ij} \} (t_i - t_j) \\ &\quad - (\sum_i t_i z_i) (NZ/A)^2 \sum (t_i - t_j)^2 \\ &\quad \times [\{ (z_i - z_j) / r_{ij} \}^2 \partial^2 V / \partial r_{ij}^2 + \{ r_{ij}'' - (z_i - z_j)'' / r_{ij}'' (\partial V / \partial r_{ij}) \}] + \dots \end{aligned} \quad (3.6)$$

* Only in the case of the Goldhaber-Teller model the internal part of the kinetic energy Hamiltonian contains neither ξ nor π .

and

$$\begin{aligned}
 (1-x)V^{(2)} &= r^2 [\pi/b, [\pi/b, V]] - \dots \\
 &= (NZ/A)^2 \sum (t_i - t_j)^2 [\{ (z_i - z_j)/r_{ij} \}^2 \partial^2 V / \partial r_{ij}^2 \\
 &\quad + \{ r_{ij}^2 - (z_i - z_j)^2 \} / r_{ij}^3 (\partial V / \partial r_{ij})] - \dots
 \end{aligned} \tag{3.7}$$

We can see that the only non-vanishing terms in the above equation are those terms which have the suffixes (i, j) related to the proton and the neutron.

Next problem is how to solve the Schrödinger equation derived from the Hamiltonian (3.4). We will formulate the quantum-mechanical dipole vibration model based on the variational principle in next section.

§ 4. Formulation of the dipole vibration model

First we assume that the wave function of the ground state of the nucleus has the separated form of variables,*

$$\Psi_0 = (1/\sqrt{A!}) \sum \varepsilon_P \psi_0(\hat{z}_P) \Phi_0(\zeta_P, \sigma_P) \phi_P, \tag{4.1}$$

which changes its sign when all the coordinates of any two particles are interchanged.

In (4.1) ϕ_P is the wave function for the isotopic spin

$$\phi_P = \alpha(1_P) \alpha(2_P) \cdots \alpha(Z_P) \beta((Z+1)_P) \cdots \beta(A_P), \tag{4.2}$$

in which α represents proton and β neutron, P stands for the operation that rearranges the set of numbers $(1, 2, \cdots A)$ to the set $(1_P, 2_P, \cdots A_P)$, \hat{z}_P and ζ_P are C-numbers which satisfy the relations,

$$\hat{z}_P \phi_P = \hat{z}_P \phi_P \quad \text{and} \quad \zeta_P \phi_P = \zeta_P \phi_P, \tag{4.3}$$

and ε_P takes the value $+1$ or -1 as the rearrangement P is even or odd. For simplicity we omit the spin variables, σ_P , in the argument of Φ_0 hereafter.

It can be shown in Appendix that the collective wave function $\psi_0(\hat{z}_P)$ and the internal wave function $\Phi_0(\zeta_P)$ have to satisfy the following integral equations in the self-consistent manner in order to minimize the expectation value of the total energy, in which the isotopic spin wave functions are eliminated and the V' is transformed into the space exchange operator P_{st} between the proton and the neutron.

$$\begin{aligned}
 &[\pi^2/2M + \{(1-x)/2\} \langle V^{(2)} \rangle_{\Phi_0} \hat{z}^2 + \cdots] \psi_0(\hat{z}) \\
 &\quad + (\Phi_0(\zeta), x \sum V_0(r_{st}) P_{st} \psi_0(\hat{z}) \Phi_0(\zeta))_{\zeta} = E_0^{(c)} \psi_0(\hat{z})
 \end{aligned} \tag{4.4a}$$

and

$$\begin{aligned}
 &[H^{(0)} + \{(1-x)/2\} V^{(2)} \langle \hat{z}^2 \rangle_{\psi_0} + \cdots] \Phi_0(\zeta) \\
 &\quad + (\psi_0(\hat{z}), x \sum V_0(r_{st}) P_{st} \psi_0(\hat{z}) \Phi_0(\zeta))_{\psi} = E_0^{(in)} \Phi_0(\zeta),
 \end{aligned} \tag{4.4b}$$

* The below discussions are not valid for the case where the total isotopic spin of the nucleus is a good quantum number.

where we have written ξ_P or ζ_P simply as $\hat{\xi}$ or $\hat{\zeta}$. It is to be noted that the $\psi_0(\hat{\xi})$ of the third term of (4.4a) can not be taken out from the integral generally.

An orthonormal set of $\psi_i(\hat{\xi})$ $\psi_j(\hat{\zeta})$'s is defined as the solutions of the following equations,

$$H_{\text{col}}\psi_i(\hat{\xi}) = E_i^{\text{col}}\psi_i(\hat{\xi}) \quad (4.5a)$$

and

$$H_{\text{in}}\psi_j(\hat{\zeta}) = E_j^{\text{in}}\psi_j(\hat{\zeta}), \quad (4.5b)$$

where

$$H_{\text{col}} = \pi^2/2M + \{(1-x)/2\} \langle V^{(2)} \rangle_{\psi_0} \hat{\xi}^2 + x \langle \sum V_0(r_{st}) P_{st} \rangle_{\psi_0} + \dots \quad (4.6a)$$

and

$$H_{\text{in}} = H^{(0)} + \{(1-x)/2\} V^{(2)} \langle \hat{\xi}^2 \rangle_{\psi_0} + x \langle \sum V_0(r_{st}) P_{st} \rangle_{\psi_0} + \dots, \quad (4.6b)$$

provided that the third terms in the right hands of (4.6)'s become the last terms in the left hands of (4.4)'s if they operate on $\psi_0(\hat{\xi})$ or $\psi_0(\hat{\zeta})$. Therefore, the total Hamiltonian of the nucleus H_N are separated into three parts,

$$H_N = H_{\text{col}} + H_{\text{in}} + H'. \quad (4.7)$$

If we put P_{st} equal to 1 tentatively, (4.4a) and (4.5a) are reduced to the Schrödinger equations of the simple harmonic oscillator. Then

$$H_{\text{col}} = \pi^2/2M + (1/2) \langle V^{(2)} \rangle_{\psi_0} \hat{\xi}^2 + \dots \quad (4.8)$$

and

$$H' = V^{(1)} \hat{\xi} + (1/2) [V^{(2)} \hat{\xi}^2 - \langle V^{(2)} \rangle_{\psi_0} \hat{\xi}^2 - V^{(2)} \langle \hat{\xi}^2 \rangle_{\psi_0}] + \dots \quad (4.9)$$

This simplified collective model will be studied in Sections 5 to 7.

§ 5. Internal coordinates

Since it is impossible to obtain the exact solution of the internal Hamiltonian, (4.6b), we will consider three conventional ways to get over the difficulties.

The simplest of them is the Hartree approximation, though we must modify it taking into account the fact that degrees of freedom are reduced. For example, the internal wave function $\psi_0(\hat{\zeta})$ can be reasonably assumed⁽⁸⁾ as follows;

$$\psi_0(\hat{\zeta}) \propto \prod_i f_i(\mathbf{x}_i) |_{\mathbf{z}=0} = \prod_i f_i(\mathbf{x}_i^{(0)}), \quad (5.1)$$

where $\prod_i f_i(\mathbf{x}_i)$ is the wave function in the ordinary Hartree approximation, which is not antisymmetrized for simplicity. The newly introduced quantity, $\mathbf{x}_i^{()}$, stands for the position vector in the coordinate system of the center of mass of neutrons (or protons) if i -th particle is a neutron (or proton). Thus the physical picture given to the wave function, (5.1), is obviously that n 's and p 's move freely in their own well. Therefore we can see that, when we use the Hartree approximation with respect to the internal coordinates,⁽⁸⁾ we get the model of "movable wells for protons and neutrons," which has been already

proposed by S. Takagi et al.⁵⁾ and S. Gallone et al.⁵⁾ independently.

It is notable that this model is compatible with the success of the $j-j$ coupling shell model for explaining the properties of medium or heavy nuclei in their ground states and we have no means to discriminate between the shell model and the model of "movable wells for protons and neutrons" by the static properties of the ground state. However, in the latter model there should exist the correlation effect of the zero-point oscillation of the dipole vibration in the ground state, which will be discussed later.

One of the other approximations is to separate also the collective coordinates corresponding to the surface vibration from the internal part of our Hamiltonian. However, we need to adopt the collective coordinates and their conjugate momenta a little different from the ones as used in S. Tomonaga et al.'s theory,^{(9),(12)} in order to make them satisfy the set of commutation relations together with $\hat{\pi}$'s and π 's of the dipole vibration. Phenomenologically this model has been studied by M. Nogami et al.¹¹⁾ and S. Gallone et al.⁵⁾

The other type of correlations, which can not be expressed by the simple displacement potential, such as α -model, d -model, will be taken into account in our theory if we can construct the corresponding forms of the internal wave function, $\psi_0(\zeta)$, though it does not seem to be an easy task.

§ 6. Estimate of the resonance frequency

In this section we estimate the magnitude of $\langle V^{(2)} \rangle_{\Phi_0}$ by the same method as T. Tamura et al.'s¹⁾ for the calculation of surface tension, assuming the uniform density and $N=Z$ for simplicity. In the form of Fourier integral,

$$\langle V^{(2)} \rangle = - \sum_{\mathbf{k}} \int f(\mathbf{k}) k_z^2 \langle e^{i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \rangle d\mathbf{k}, \quad (6.1)$$

where $f(\mathbf{k})$ is the Fourier coefficient of $V_0(r)$. The assumption of uniform density within the spherical region with radius $R=r_0 A^{1/3}$ leads to

$$\begin{aligned} \langle V^{(2)} \rangle = & - (3\pi/r_0^4) \int_0^\infty f(k) \{ R^2/2 + (R^2/2) \cos 2kR + 1/2k^2 - \cos(2kR)/2k^2 \\ & - (R/k) \sin(2kR) \} dk. \end{aligned} \quad (6.2)$$

In the case of Yukawa well, $V_0(r) = V_0 e^{-\mu r}/\mu r$ and $f(k) = V_0/2\pi^2 \mu (k^2 + \mu^2)$,

$$\langle V^{(2)} \rangle \cong - (3V_0 A^{1/3}/8\mu^2 r_0^4) (1 - 1/\mu^2 R^2). \quad (6.3)$$

In the case of Gaussian type interaction, $V_0(r) = V_0 e^{-\beta r^2}$ and $f(k) = (V_0/8(\beta\pi)^{3/2}) e^{-k^2/4\beta}$,

$$\langle V^{(2)} \rangle = - (3V_0 A^{2/3}/16\beta r_0^4) (1 - 1/2\beta R^2). \quad (6.4)$$

The parameters contained in the two-body interactions are determined from the analysis of the nucleon-nucleon scattering,¹⁴⁾ where Serber's exchange character, $\alpha=1/2$ is taken. For Yukawa well $V_0 = -51.0$ Mev, $\mu = 1/1.125 \times 10^{13} \text{ cm}^{-1}$ and for Gaussian well $V_0 = -59.6$ Mev and $\sqrt{\beta} = 0.644 \times 10^{13} \text{ cm}^{-1}$.

We obtain the value of the resonance energy of the simplified collective model (4.8), $\hbar\nu = \hbar\sqrt{\langle V^{(2)} \rangle / M}$, by inserting the numerical values of $\langle V^{(2)} \rangle$. It is interesting to notice that (6.3) and (6.4) have the same A -dependence as the Goldhaber-Teller's original prediction except the small correction terms. The results are shown in Fig. 2, together

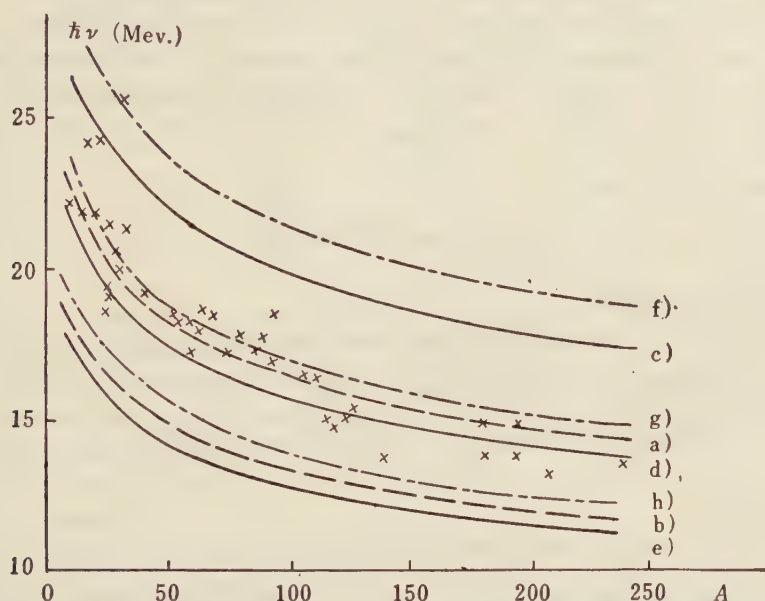


Fig. 2. Resonance energy $\hbar\nu$: The crosses give the experimental data.¹⁵⁾ The curves, a) and b), stand for the Fujii-Takagi's results⁵⁾ using the uniform density model for $r_0 = 1.35$ and 1.50×10^{-13} cm and the square well type interaction. c), d) and e) represent the uniform density model using the Yukawa well potential $V_0 e^{-\mu r} / \mu r$ ($V_0 = -51.0$ Mev. and $\mu = 1/1.125 \times 10^{13}$ cm⁻¹) for $r_0 = 1.2, 1.35$ and 1.50×10^{-13} cm respectively. f), g) and h) are the corresponding curves for the Gaussian well potential $V_0 e^{-\beta r^2}$ ($V_0 = -59.6$ Mev. and $\sqrt{\beta} = 0.644 \times 10^{13}$ cm⁻¹).

with the Fujii-Takagi's results⁵⁾ using the square well type interaction. All of them agree reasonably with the experimental data,¹⁵⁾ if we choose the appropriate nuclear radii. However, in Section 8 the additional effects of $n-p$ exchange forces are discussed and it will become clear that the numerical value of the radius parameter r_0 which gives the best fit to the experimental data has almost no quantitative meaning.

§ 7. Interaction terms

The couplings between the dipole vibration and the internal coordinates come from the interaction Hamiltonian, H' , which has been called the internal friction phenomenologically and consists of $H^{(1)} \dot{\xi}$, etc., quantum-mechanically. Though the diagonal parts of $H^{(1)}$ should vanish as we expect from the classical analogy, the non-diagonal parts of $H^{(1)}$ will be very important to determine the fine structure of the giant resonance. When we

study the coupling effect, it can be known how wide the peak of the giant resonance is or how much the true ground state deviates from (4.1). If the effect is so large that the true ground state wave function deviates remarkably from (4.1), our collective model has little significance. However, we know in this section that it is difficult to obtain a reliable estimate to give a criterion whether the dipole vibration model is valid or not.

First we will develop the order of magnitude theory using the method by which E. P. Wigner et al.⁽¹²⁾ discussed the width of the resonance in the neutron scattering. The function X_λ of the compound nucleus, which satisfy the wave equation,

$$H_N X_\lambda = E_\lambda X_\lambda, \quad (7.1)$$

are decomposed in the form

$$X_\lambda = \sum_{ij} C_{\lambda; ij} \psi_i(\xi) \phi_j(\zeta), \quad (7.2)$$

where $\psi_i(\xi) \phi_j(\zeta)$ satisfies

$$(H_N - H') \psi_i(\xi) \phi_j(\zeta) = E_{ij} \psi_i(\xi) \phi_j(\zeta). \quad (7.3)$$

According to the reference 13, it can be proved that

$$M^{(v)} \equiv \sum_\lambda (E_\lambda - E_{ij})^2 C_{\lambda; ij}^2 = (\psi_i \phi_j, (H_N - E_{ij})^2 \psi_i \phi_j). \quad (7.4)$$

Then we obtain

$$\sum \{E_\lambda - (E_{ij} + M_{ij}^{(1)})\}^2 C_{\lambda; ij}^2 = M_{ij}^{(2)} - (M_{ij}^{(1)})^2. \quad (7.5)$$

Therefore we can see that the quantity $M_{(0)}^{(2)} - (M_{(0)}^{(1)})^2$ should provide a measure of deviation of the ground state from the state $\psi_0 \phi_0$. On the other hand, if the dipole vibration model is valid and $C_{(0)}^2$ can be regarded approximately as 1, $M_{(0)}^{(2)} - (M_{(0)}^{(1)})^2$ should provide* a measure of the square of the widths Γ , since by E1 interaction the $\psi_0 \phi_0$ can transit only to the $\psi_1 \phi_0$.

The calculation of $M_{(0)}^{(2)}$ for the simplified interaction (4.9) can be easily carried out** if the surface correction is neglected.⁽¹⁾ Using the Gaussian well, $V(r) = V_0 e^{-\beta r^2}$,

$$\begin{aligned} M_{10}^{(2)} - (M_{10}^{(1)})^2 &= (9 \hbar^2 V_0^2 / 32 m r_0^3 \hbar \nu) \sqrt{2\pi/\beta} + (27 \hbar^2 V_0^2 / 32 A m \hbar \nu) \\ &\times r_0^7 \beta^3 \sqrt{\beta \pi / 2} A^{2/3} + (27 \hbar^4 V_0^2 / 32 m^2 (\hbar \nu)^2 A r_0^3) \sqrt{2\beta \pi} + \dots, \end{aligned} \quad (7.6)$$

in which the first and second terms in the right hand come from $H^{(1)}$ and the third from $H^{(2)}$. The contribution from $H^{(1)}$ is predominantly large and the others are less than 8% even at $A=20$. The parameters of the nuclear forces are assumed to be the same as appeared in Section 6 and $r_0=1.37$ and 1.50. The numerical results are shown in Table I, together with the half of the empirical width $\Gamma/2$.

* It may happen that the distribution of $C_{\lambda; 10}^2$ has a long tail in the high energy region because of the effect of H' and the calculation of the second order moment gives an overestimate for the half-width.

** In our case, since the interactions exist only between neutrons and protons, it is almost unnecessary to consider the correlations due to the Pauli principle.

Table I. Calculated values of $\sqrt{M_{10}^{(2)} - (M_{10}^{(1)})^2}$, the square root of the second moment of the interaction Hamiltonian: We assume the uniform density model and the nuclear potential, $V_N = \sum_{st} (V_0/2) (1 + P_{st}) \exp(-\beta r_{st}^2)$, in which $V_0 = -59.6$ Mev. and $\sqrt{\beta} = 0.644 \times 10^{13} \text{ cm}^{-1}$.

| $\begin{array}{c} \backslash \\ A \end{array} \begin{array}{c} r_0 \\ \end{array}$ | | $\sqrt{M_{10}^{(2)} - (M_{10}^{(1)})^2} \text{ (Mev.)}$ | | | | $(\Gamma/2)_{\text{obs.}} \text{ (Mev.)}$ |
|--|--|---|------|------------|------|---|
| | | $P_{st}=1$ | | $P_{st}=0$ | | |
| | | 1.37 | 1.50 | 1.37 | 1.50 | |
| 20 | | 72.1 | 59.9 | 36.0 | 30.0 | ~ 4.5 |
| 50 | | 74.2 | 61.2 | 37.1 | 30.6 | |
| 100 | | 75.4 | 62.5 | 37.7 | 31.3 | ~ 4.0 |
| 150 | | 76.3 | 63.5 | 38.2 | 31.7 | ~ 3.5 |
| 200 | | 77.1 | 64.3 | 38.5 | 32.1 | |

Apparently our results fail to explain the absolute values of the experimental half-widths of the giant resonance. Now we have two possibilities: (1) The effect of the interactions is so large that our orthonormal set of the dipole vibration model is not useful. (2) We have obtained too large estimate for the widths because of the assumed uniform density wave function being wrong, as seen from the fact that we can not obtain the finite results for Yukawa well. If we assume the existence of the hard core, the results will be greatly improved.*

Also the theoretical half-widths increase a little with A in contrast with the observed ones. However, it may be necessary to take into account the detailed structure of the individual nucleus to discuss the half-widths of giant resonance, since the fluctuation of the observed data is remarkable.

Next, we will study phenomenologically the rough relation between the impurity of the ground state and the wide width of the giant resonance. According to the above results we can reasonably assume that the main contribution comes from $H^{(1)}\xi$ in H' . Then

$$M_{10}^{(2)} - (M_{10}^{(1)})^2 \cong 3 (M_{00}^{(2)} - (M_{00}^{(1)})^2). \quad (7.7)$$

If we put the left hand side equal to the observed $(I'/2)^2 \approx 16 \text{ Mev}^2$ phenomenologically, we get $M_{00}^{(2)} - (M_{00}^{(1)})^2 \approx 16/3 \text{ Mev}^2$. Since our choice of H' in Section 4 leads necessarily to $(\psi_i \phi_j, H' \phi_0 \phi_0) = 0$ for either i or $j=0$, except $i=j=0$. Therefore

$$\begin{aligned} |C_{0;00}|^2 &\cong 1 - \sum |(\psi_i \phi_j, H' \phi_0 \phi_0) / (E_{00} - E_{ij})|^2 \\ &\gtrsim 1 - 3 (M_{00}^{(2)} - (M_{00}^{(1)})^2) / (\hbar\nu)^2 \sim 13/14 \end{aligned} \quad (7.8)$$

assuming the validity of the first order perturbation theory and $\hbar\nu \sim 15 \text{ Mev}$. The factor 3 is added, since our space is three-dimensional.

The equation (7.8) tells us that the wide width of the giant resonance, $I' \sim 8 \text{ Mev}$,

* It is interesting that our situation is quite similar to the intermediate coupling optical model by E. P. Wigner et al.¹³⁾.

can be explained* if there exists the configuration mixing of the probability $1/14$. The ground state may be regarded approximately as $\Psi_0 \sim \psi_0 \Phi_0$, if the dipole vibration model is valid.

§ 8. The effect of neutron-proton exchange forces

It is tried in this section to survey the effects of neutron-proton exchange forces and clarify the physical meaning of Levinger-Bethe's general argument¹⁰⁾ using the sum rules. We can construct a new collective model taking into account the exchange effect, which we will call the Modified Collective Model (M.C.M.).

The fundamental equation of M.C.M. is the equation, (4.5a);

$$(\pi^2/2M + ((1-x)/2) \langle V^{(2)} \rangle \hat{\xi}^2 + \dots) \psi_i(\hat{\xi}) + (\Phi_0(\xi), x \sum V_0(r_{st}) P_{st} \psi_i(\hat{\xi}) \Phi_0(\xi)) \tau_i = E_i^{\text{col}} \psi_i(\hat{\xi}). \quad (4.5a).$$

It is shown that the integral with respect to all the internal coordinates in the left hand side of (4.5a) has the velocity-dependent character, namely the different effects for the different ψ_i 's.

If we confine ourselves to the ground state ($i=0$) and the first collective excited state ($i=1$),** we can rewrite (4.5a), neglecting the higher order terms of $\hat{\xi}$,

$$\{\pi^2/2M + ((1-x)/2) \langle V^{(2)} \rangle \hat{\xi}^2 + xS_i + (x/2) K_i \hat{\xi}^2 + \dots\} \psi_i(\hat{\xi}) = E_i^{\text{col}} \psi_i(\hat{\xi}) \quad (8.1)$$

where $x/2 \cdot K_i \hat{\xi}^2$ stands for the contribution from the neutron-proton exchange forces to the resonance energy and xS_i the additional level shift for the state ψ_i .

The aim of this section is to obtain the estimates for K_i and S_i . First let us study the effect of P_{st} to $\psi_0(\hat{\xi})$. We can assume that

$$\psi_0(\hat{\xi}) = N_0 e^{-1/2 \cdot \alpha^2 \hat{\xi}^2}. \quad (8.2)$$

Putting $N=Z$ and using the relation,

$$z_s = z_s^{(0)} - (1/2) \tau_{ss} \hat{\xi}, \quad (8.3) ***$$

then

$$P_{st} \hat{\xi} = \hat{\xi} - u_{st} \quad (8.4)$$

where

$$u_{st} = (4/A) \{ (z_s^{(0)} - z_t^{(0)}) + \hat{\xi} \}. \quad (8.5)$$

Thus

* By the same reason as the first footnote of this section the factor 3 of (7.7) can not be taken seriously. The higher order perturbation terms may also decrease the value of $|C_{0;00}|^2$. Thus the numerical value $1/14$ has little quantitative meaning.

** For the larger suffix i than one it is questionable that (4.5a) can be rewritten at once in the form of (8.1), since the states with lower suffixes may also appear in (8.1).

*** It is the characteristic of Goldhaber-Teller model that there exists no term having the higher power than one.

$$P_{st}\psi_0(\hat{\xi}) = N_0 e^{-(1/2)\alpha^2(\hat{\xi}-u_{st})^2} \\ \cong \psi_0(\hat{\xi}) e^{-(8/A^2)\alpha^2(z_s^{(0)}-z_t^{(0)})^2} \{1 + D_1^{(st)}\hat{\xi} + (1/2)D_2^{(st)}\hat{\xi}^2 + \dots\}, \quad (8.6)$$

where

$$D_1^{(st)} = (4\alpha^2/A) (z_s^{(0)} - z_t^{(0)}) \quad (8.6a)$$

$$D_2^{(st)} = 8\alpha^2/A + (16\alpha^4/A^2) (z_s^{(0)} - z_t^{(0)})^2. \quad (8.6b)$$

Similarly we assume for the first excited state as

$$\psi_1(\hat{\xi}) = N_1 2\alpha' \hat{\xi} e^{-(1/2)\alpha'^2 \hat{\xi}^2}. \quad (8.7)$$

Then

$$P_{st}\psi_1(\hat{\xi}) = \psi_1(\hat{\xi}) e^{-(8/A^2)\alpha'^2(z_s^{(0)}-z_t^{(0)})^2} \{1 - 4/A + D_0^{(st)'} + D_1^{(st)'}\hat{\xi} \\ + (1/2)D_2^{(st)'}\hat{\xi}^2 + \dots\} - N_1 2\alpha' (4/A) (z_s^{(0)} - z_t^{(0)}) e^{-(8/A^2)\alpha'^2(z_s^{(0)}-z_t^{(0)})^2} - (\alpha'^2/2)\hat{\xi}^2, \quad (8.8)$$

where

$$D_0^{(st)'} = -(16\alpha'^2/A^2) (z_s^{(0)} - z_t^{(0)})^2 \quad (8.9)$$

and $D_1^{(st)'}$ and $D_2^{(st)'}$ are obtained by inserting α' instead of α in (8.6a) and (8.6b) respectively.

Also we can expand $P_{st}\Phi_0(\zeta)$ and $V_0(r_{st})$ into power series of $\hat{\xi}$:

$$P_{st}\Phi_0(\zeta) = \Phi_0^{(st)}(\zeta) + \Phi_0^{(st)'}(\zeta)\hat{\xi} + (1/2)\Phi_0^{(st)''}(\zeta)\hat{\xi}^2 + \dots \quad (8.10)$$

$$V_0(r_{st}) = V_{st}^{(0)} + V_{st}^{(1)}\hat{\xi} + (1/2)V_{st}^{(2)}\hat{\xi}^2 + \dots \quad (8.11)$$

The contribution from $\Phi_0^{(st)'}(\zeta)$ can be reasonably neglected* for sufficiently large A . Substituting these equations into (4.5a) and comparing with (8.1), we obtain

$$\mathcal{S}_1 - \mathcal{S}_0 = -(16\alpha'^2/A^2) (\Phi_0, \sum V_{st}^{(0)} (z_s^{(0)} - z_t^{(0)})^2 \Phi_0^{(st)}) \\ - (4/A) [(\Phi_0, \sum V_{st}^{(0)} \Phi_0^{(st)}) + (\Phi_0, \sum V_{st}^{(1)} (z_s^{(0)} - z_t^{(0)}) \Phi_0^{(st)})] + \dots, \quad (8.12)$$

$$K_0 = (\Phi_0, \sum V_{st}^{(2)} \Phi_0^{(st)}) + (8\alpha^2/A) [(\Phi_0, \sum V_{st}^{(0)} \Phi_0^{(st)}) + (\Phi_0, \sum V_{st}^{(1)} (z_s^{(0)} - z_t^{(0)}) \Phi_0^{(st)})] \\ + [(16\alpha^4/A^2) (\Phi_0, \sum V_{st}^{(0)} (z_s^{(0)} - z_t^{(0)})^2 \Phi_0^{(st)}) + (\Phi_0, \sum V_{st}^{(0)} \Phi_0^{(st)''})] + \dots \quad (8.13a)$$

and

$$K_1 = K_0' + \dots, \quad (8.13b)$$

where K_0' involves $(\alpha')'$'s instead of α 's in K_0 .

Since the difference between K_1 and K_0' is approximately independent of A , the frequencies of the dipole vibrations in the ground and the first collective excited states are practically the same for large A . In order to determine the resonance frequency $\hbar\nu$ theoretically, we can use the relations,

$$\hbar\nu = \hbar \sqrt{\{(1-x)\langle V^{(2)} \rangle + xK_0'\}/M} \quad \text{and} \quad \alpha^2 = M \cdot \hbar\nu/\hbar^2. \quad (8.14)$$

* We can check it by using the Fermi gas model.

Therefore we obtain $\alpha=\alpha'$. It is noticed that each of the second to the fifth terms of (8·13a) has probably a large absolute value compared with the first term. Unless they almost cancel each other,* the discussion of Section 6 will lose its value entirely. It seems difficult to derive a definite conclusion about it.

The observed level spacing $(\hbar\nu)_{\text{obs.}}$ between the ground state and the first collective excited state is the sum of the resonance energy of the dipole vibration and the additional level shift $x(S_1-S_0)$. The first term in the right hand side of (8·12) can be easily evaluated by making use of the calculated value⁽¹⁶⁾ in the Fermi gas model to a first approximation :

$$\begin{aligned} & -(16\alpha^2/A^2) (\Phi_0, \sum V_{st}^{(0)} (z_s^{(0)} - z_t^{(0)})^2 \Phi_0^{(st)}) \\ & \cong -(16\alpha^2/3A^2) (\Psi_0, \sum V_0(r_{st}) r_{st}^2 P_{st} \Psi_0) \\ & \cong 0.8 \hbar\nu. \end{aligned} \quad (8 \cdot 15)$$

On the other hand the second and the third terms of (8·12) can be rewritten by the partial integration as follows ;

$$\begin{aligned} \delta' & \equiv -(4/A) [(\Phi_0, \sum V_{st}^{(0)} \Phi_0^{(st)}) + (\Phi_0, \sum V_{st}^{(1)} (z_s^{(0)} - z_t^{(0)}) \Phi_0^{(st)})] \\ & \cong (2/A) \sum_{st} \int V(r_{st}) (z_s - z_t) (\partial/\partial z_s - \partial/\partial z_t) (\Psi_0^* P_{st} \Psi_0) d\tau. \end{aligned} \quad (8 \cdot 16)$$

If we ignore the exchange operator P_{st} , the right hand side presumably vanishes. Thus it is a quite delicate problem to evaluate the magnitude. However, for purposes of illustration, let us adopt the same model⁽¹⁶⁾ as are necessary in order to obtain the value 0.8 of (8·15). Assuming the square well potential with depth V_0 and range b ,

$$\delta' = 18V_0/(kr_0)^3 \left[(1/2) \int_0^{kb} (\sin x/x) dx - kb (j_0^2(kb) + j_1^2(kb)) \right], \quad (8 \cdot 17)$$

where k is the maximum wave number for nucleons in the nucleus. The quantity in the large parentheses of (8·17) is quite sensitive to the assumed value of kb ; 1/200, 1/20 and 1/5 for $kb=1.0, 1.8$ and 2.5 respectively. The parameters of the nuclear potential in the eq. (22) of reference 16 and $r_0=1.50$ lead to the results, $\delta'=13.5$ Mev for triplet and 15.8 Mev for singlet, which are too large to be compatible with the observed facts. However, the numerical results do not seem to be very reliable because of the involved uncertainties.

In the result

$$(\hbar\nu)_{\text{obs.}} = (1 + x\delta) \hbar\nu, \quad (8 \cdot 18)$$

where

$$\delta \approx 0.8 + \delta'/\hbar\nu. \quad (8 \cdot 19)$$

* We can prove that the third and fifth terms of (8·13a) are positive and the second and fourth terms are negative in the Fermi gas model, though the total contribution may not vanish exactly.

Since δ' is probably positive, the resonance energy of dipole vibration $\hbar\nu$ is expected to be somewhat smaller* than the observed one if there exist Majorana forces. This situation is shown in Figure 3 schematically.

We can easily obtain the sum rules in M.C.M. using (8·18) :

$$\begin{aligned} \int \sigma dW &\cong (4\pi^2 e^2 / \hbar c) (NZ/A)^2 (\hbar\nu)_{\text{obs.}} \\ &\times (\Psi_0, \xi^2 \Psi_0) \\ &\cong (2\pi^2 e^2 \hbar / mc) (NZ/A) (1 + x\delta), \end{aligned} \quad (8 \cdot 20)$$

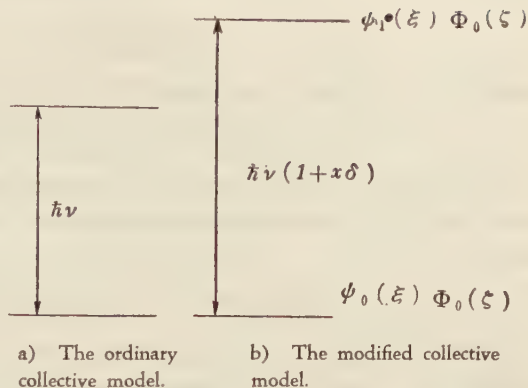


Fig. 3. Schematic picture of the additional effect of the n - p exchange forces.

which agrees with Levinger-Bethe's sum rule¹⁶⁾ if δ' is small.

$$\begin{aligned} \int (\sigma/W) dW &= 4\pi^2 (e^2 / \hbar c) (NZ/A)^2 (\Psi_0, \xi^2 \Psi_0) \\ &\cong (2\pi^2 e^2 \hbar / mc) (NZ/A) (1 + x\delta) / (\hbar\nu)_{\text{obs.}} \end{aligned} \quad (8 \cdot 21)$$

and

$$W_H = \int \sigma dW / \int (\sigma/W) dW \cong (\hbar\nu)_{\text{obs.}}, \quad (8 \cdot 22)$$

which means that the allowed levels are concentrated over the energy region near to $(\hbar\nu)_{\text{obs.}}$. Because of the interaction Hamiltonian H' , Ψ_0 must deviate somewhat from the pure state $\psi_0(\xi) \Phi_0(\zeta)$, but the deviation will not be large if the phenomenological estimate of Section 7 is valid.

§ 9. Correlation effect of the dipole vibration in the nuclear ground state

The success of the Goldhaber-Teller model for explaining the A -dependence of the photon energy corresponding to the maximum absorption cross section, if we may say so, suggests simultaneously that the deviation of the true wave function from $\psi_0(\xi) \Phi_0(\zeta)$ is small. The aim of this section is to study the possibility for the validity of the dipole vibration model.

First we will discuss the relative superiority among the ordinary independent particle model and the model of "the movable wells for neutrons and protons" from the viewpoint of the variational principle. The ground state of the Goldhaber-Teller model, Ψ_0^{GT} , is assumed for simplicity as⁸⁾

* The discussion of this section can be reformulated also as follows:

$$(\psi_1 \Phi_0, H_N \psi_1 \Phi_0) - (\psi_0 \Phi_0, H_N \psi_0 \Phi_0) = \hbar\nu + x(S_1 - S_0) = \hbar\nu(1 + x\delta).$$

$$\Psi_0^{GT} \cong \phi_0(\hat{\zeta}) \Phi_0(\zeta) \propto \phi_0(\hat{\zeta}) [\Psi_0^H]_{\zeta=0}, \quad (9.1)$$

where Ψ_0^H is the wave function in the ordinary independent particle model.

$$(\Psi_0^{GT}, H_N \Psi_0^{GT}) \cong (3/2) \hbar \nu - (\Psi_0^H, (\pi^2/2M) \Psi_0^H) + (\Psi_0^H, H_N \Psi_0^H), \quad (9.2)$$

where we have neglected the effect of higher power terms than $\hat{\zeta}^2$ as well as the additional effect due to the reduction of the degrees of freedom. If we neglect the interference terms,

$$(1/2M) (\Psi_0^H, \pi^2 \Psi_0^H) \cong -(\hbar^2/2mA) (\Psi_0^H, \sum_i \Delta_i \Psi_0^H) = \langle T \rangle / A. \quad (9.3)$$

Therefore the relative superiority between both models can be determined if it is known whether the average kinetic energy per nucleon is larger than the three dimensional zero-point oscillation energy, $(3/2) \hbar \nu$. If we put the Fermi gas model, $\langle T \rangle / A \approx 14.3$ Mev for $r_0 = 1.37$ and the larger for the smaller radius. Though $(\hbar \nu)_{\text{obs.}} \approx 15$ Mev for medium nuclei, $(\hbar \nu)$ is probably smaller by about 30% than $(\hbar \nu)_{\text{obs.}}$ because of the exchange effect. Thus $(3/2) \hbar \nu \sim 15.7$ Mev and it is possible that the dipole vibration model has comparable value compared with the independent particle model as a model for the ground state of the sufficiently large nuclei.

Next let us study the direct experimental evidence of the correlation effect in the nuclear ground state. It is already known⁽¹⁾ that the experimental harmonic mean energy W_H for photon absorption is nearly equal to the resonance frequency $(\hbar \nu)_{\text{obs.}}$. According to the conventional way⁽¹⁶⁾ we can obtain the absorption area of photon from W_H by the assumption of no correlation. It is usually understood^{(16),(17)} that the comparison of $\langle r^2 \rangle_{00}$

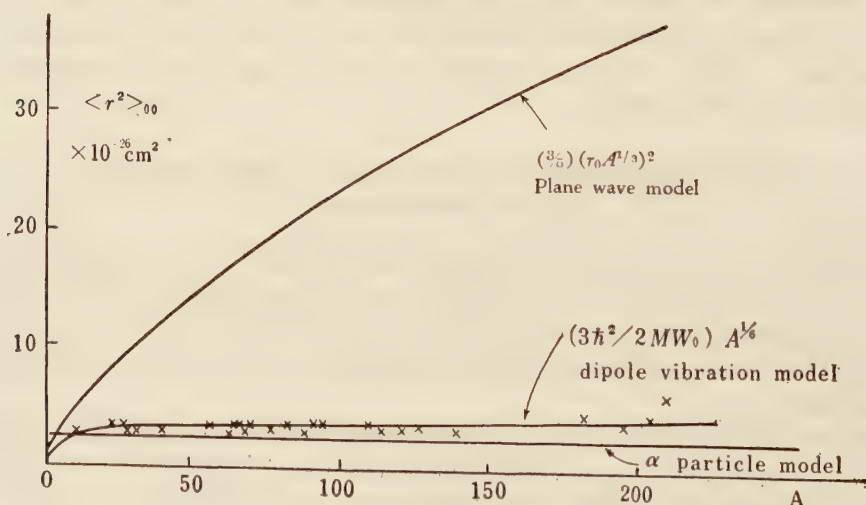


Fig. 4. The crosses give the values of $\langle r^2 \rangle_{00}$ which are obtained⁽¹⁷⁾ from the empirical W_H 's assuming $\chi=0$. The strongly ascending curve represents the calculated values of $\langle r^2 \rangle_{00}$ obtained from the plane wave model of the nucleus and the straight line parallel to the abscissa the value of $(3/5) (r_0 A^{1/3})^2$ for $A=4$, for which we assume tentatively $r_0 = 1.35 \times 10^{-13}$ cm. The other curve stands for the dipole vibration model with $\chi=0$.

with the experimental data strongly favours the alpha-particle model as shown in Figure 4. However, we have not the definite experimental evidences for the validity of the alpha-particle model except for the light nuclei. The situation that the uniform absorption over the total area of the nucleus can not explain the experimental $\langle r^2 \rangle_{00}$ as shown in Fig. 4, needs modification by taking into account the correlation due to the Pauli principle between like nucleons, for which quite elaborate analyses have been carried out by J. S. Levinger et al.,³⁾ the neutron-proton correlations being neglected.

In the dipole vibration model the quantity corresponding to the $\langle r^2 \rangle_{00}$ can be easily calculated using (8.21) and (8.22). Our results for $\kappa=0$ are compared with the empirical data of $\langle r^2 \rangle_{00}$ in Fig. 4, which J. Goldemberg et al.¹⁷⁾ obtained from the empirical W_n 's assuming $\kappa=0$. In the numerical calculation, the empirical relation $\hbar\nu \approx W_0 A^{-1/6} \sim 36 A^{-1/6}$ Mev is used, which has the A -dependence predicted by Goldhaber-Teller and is not inconsistent with the experimental data. The agreement with experiment is very good and the correlation effect of the dipole vibration in the ground state may give an alternative explanation instead of the alpha-particle model at least for the medium or heavy nuclei.

However, there are significant exchange forces in actual nuclei, hence to put $\kappa=0$ is wrong and the experimental values $\langle r^2 \rangle_{00}$ in Fig. 4 must be somewhat increased. Simultaneously $\hbar\nu$ must be taken to be small by the same factor compared with the observed value $(\hbar\nu)_{\text{obs}}$, as seen from (8.22). Thus the above conclusion needs no change.

Recently it has become clear¹²⁾ that the E_1 gamma-ray transitions of low energies are quite suppressed in the heavy nuclei. It seems to suggest strongly the existence of correlations between neutrons and protons in the nuclear ground states.

§ 10. Relation to the other theories and discussions

In this paper Goldhaber-Teller's model, the simplest type of collective model, was reformulated as M.C.M., taking into account the additional effects of $n-p$ exchange forces. It was emphasized that the existence of the dipole vibration is compatible with the success of the $j-j$ coupling shell model in explaining the static properties of medium or heavy nuclei. Therefore the most interesting problem is which of the ordinary independent particle model and the model of "movable wells for neutrons and protons" seems to be the better model. From the viewpoint of the variation principle the relative superiority was discussed in Section 9 and the conclusion suggests that both models may have comparable values because of the effect of exchange forces.

Phenomenologically the two points favour the dipole vibration model: One of them is that the smooth A -dependence of the photon energies corresponding to the maximum absorption cross section can be naturally explained as first pointed out by Goldhaber-Teller¹⁾ and the absolute values of them are consistent⁵⁾ with the data of the two-body nuclear forces, though the effects of $n-p$ exchange forces are not clear as remarked in Section 8. All the other theories of the photoreaction do not seem to be very successful as to this point. Another advantage of the dipole vibration model is that the systematics

of the harmonic mean energies of the photon absorption cross section can be quite easily understood¹⁾ by the idea that the states to which the nucleus can transit by E_1 absorption from the ground state, are concentrated over the certain narrow energy range. The observed A -dependence of the harmonic mean energies is entirely different from the prediction of the independent particle model, in which no correlation between nucleons is considered. The introduction of the Pauli principle¹⁾ can give a better explanation, but it does not seem to be completely convincing. Whether the model of movable wells is excellent or not with respect to the other various points, for which the shell model is successful, are left for further investigation.

The direct evaluation of the second moment of the interaction Hamiltonian fails to explain the observed widths of the giant resonance as shown in Section 7. However, S. Takagi et al.⁵⁾ and S. Gallone et al.⁵⁾ assumed the interaction* between the collective motion and the internal motion phenomenologically as

$$H' = -\sum_i (\tau_{zi}/2) \hat{\xi} \partial U(r_i) / \partial z_i + \dots \quad (10.1)$$

($N=Z$)

which can be easily derived from $H^{(1)}$ in our formalism, if we assume $V = \sum_i U(r_i)$ instead of (3.3) and insert it into (3.6). The fact⁷⁾ that the phenomenological interaction can give an estimate having a right order of magnitude for the absorption widths seems to suggest the inadequacy of the assumed form of internal wave function for the calculation of the second moment, though the relation between the second moment and the absorption width is also not clear-cut.

The other well-known collective model of photoreaction is the Steinwedel-Jensen model,²⁾ which treats the relative motion of protons and neutrons in the nuclear matter hydrodynamically imposing the two conditions that the total density is constant and the surface is fixed. The coordinate $\hat{\xi}$ corresponding to the Steinwedel-Jensen model is

$$\hat{\xi} \sim \sum_i t_i j_1(kr_i) Y_{10}(\varrho_i) \quad (10.2)$$

instead of $\hat{\xi} \sim \sum_i t_i r_i Y_{10}(\varrho_i)$ in the Goldhaber-Teller model. Though we can formulate it quantum-mechanically in a similar manner, we meet with several additional complexities as compared with the Goldhaber-Teller model; (1) $[\pi, \hat{\xi}] = -i\hbar$ can be satisfied only approximately, for which the applicability of "the law of large number"⁶⁾ is questionable. (2) Electric dipole interaction contains not only $\hat{\xi}$ but ζ . (3) The coupling terms of the kinetic energy involving $\hat{\xi}$ appear. However, it seems that the best collective model should be somewhat modified¹⁰⁾ from the Goldhaber-Teller model to the Steinwedel-Jensen model. Though the empirical symmetrical energy gives somewhat small resonance frequency,²⁾ the difficulty will be well settled by the same reason as shown in Section 8.

J. M. Araújo³⁾ considered also the compressibility of the nuclear matter adding to the above Steinwedel-Jensen model, which can also be formulated similarly as in this paper. But in this case the situation becomes far more entangled.

* The surface interaction adopted by Bohr-Mottelson¹⁸⁾ just corresponds to it in our case.

M. Nogami et al.⁽¹¹⁾ and S. Gallone et al.⁽⁵⁾ treated the coupling between the dipole vibration and the surface vibration. K. Okamoto⁽¹⁰⁾ discussed the relation between the nuclear deformation and the absorption width of photon. The extension of our theory to these directions is straightforward in principle as already referred to in Section 5.

Previously D. C. Peaslee⁽¹¹⁾ has tried to connect the classical picture of the dipole vibration with the configurational mixing in the shell model quantum-mechanically and explain why the maximum cross-section of the giant resonance is so large. In our theory the excited state, to which the electric dipole transition from the ground state, $\Psi_0 \approx \phi_0(\xi) \cdot \Phi_0(\zeta)$, is allowed, is $\Psi_G \equiv \phi_1(\xi) \Phi_0(\zeta)$, although Ψ_G may be shared over the broad energy region because of the coupling terms. Omitting the constant factor,

$$\Psi_G \sim \xi \Psi_0. \quad (10 \cdot 3)$$

This equation clarifies the correspondence to the configuration mixing such as discussed by D. C. Peaslee,⁽¹¹⁾ that the excited state is a superposition of the jumped states with the different phases for neutrons and protons from the viewpoint of the independent particle model.

This work was done chiefly during the term of the seminar of "many-body problem" held at Research Institute for Fundamental Physics. The author wishes to express his sincere thanks for the stimulating discussions of Professors S. Tomonaga, M. Nogami, S. Hayakawa and Dr. T. Tamura and many other members who attended the seminar. The author is also much indebted to Professors M. Kobayasi, S. Takagi, Y. Fujimoto and many other members of Kyoto University and Professors T. Yamanouchi, S. Nakamura, Messrs. A. Sugie, T. Kikuta and M. Yamada of Tokyo University for many valuable suggestions and discussions.

Appendix

For simplicity we consider the case of two independent variables, ξ and ζ .

$$H = H_1(\xi) + H_2(\zeta) + H_{12}, \quad (A \cdot 1)$$

Taking the trial function,

$$\Psi_0 = \phi_0(\xi) \Phi_0(\zeta), \quad (A \cdot 2)$$

we want to minimize the expectation value of H with respect to Ψ_0 :

$$(\delta \Psi_0, H \Psi_0) = (\delta \phi_0(\xi) \cdot \Phi_0(\zeta), H \phi_0(\xi) \Phi_0(\zeta)) + (\delta \Phi_0(\zeta) \cdot \phi_0(\xi), H \phi_0(\xi) \Phi_0(\zeta)) \quad (A \cdot 3)$$

and

$$(\delta \Psi_0, \Psi_0) = (\delta \phi_0(\xi), \phi_0(\xi)) + (\delta \Phi_0(\zeta), \Phi_0(\zeta)) \quad (A \cdot 4)$$

Therefore $\delta \{(\Psi_0, H \Psi_0) / (\Psi_0, \Psi_0)\} = 0$ is equivalent to

$$\{H_1(\xi) + \langle H_2(\zeta) \rangle_{\Phi_0}\} \phi_0(\xi) + (\Phi_0(\zeta), H_{12} \phi_0(\xi) \Phi_0(\zeta))_{\xi} = E_0 \phi_0(\xi) \quad (A \cdot 5a)$$

and

$$\{H_2(\zeta) + \langle H_1(\hat{\zeta}) \rangle_{\psi_0}\} \Phi_0(\zeta) + (\psi_0(\hat{\zeta}), H_{12}\psi_0(\hat{\zeta}) \Phi_0(\zeta))_{\zeta} = E_0 \Phi_0(\zeta). \quad (\text{A} \cdot 5b)$$

Thus

$$H_1(\hat{\zeta})\psi_0(\hat{\zeta}) + (\Phi_0(\zeta), H_{12}\psi_0(\hat{\zeta}) \Phi_0(\zeta))_{\zeta} = \{E_0 - \langle H_1(\zeta) \rangle_{\psi_0}\} \psi_0(\hat{\zeta}) \quad (\text{A} \cdot 6a)$$

and

$$H_2(\zeta) \Phi_0(\zeta) + (\psi_0(\hat{\zeta}), H_{12}\psi_0(\hat{\zeta}) \Phi_0(\zeta))_{\zeta} = \{E_0 - \langle H_1(\hat{\zeta}) \rangle_{\psi_0}\} \Phi_0(\zeta), \quad (\text{A} \cdot 6b)$$

which are so-called Hartree equations.

It is straightforward to apply the above theorem to our problem: (A·1) corresponds to (3·4), and instead of (A·2) we assume the trial function (4·1). Then

$$\begin{aligned} (\Psi_0, H\Psi_0) &= (\psi_0(\hat{\zeta}_P) \Phi_0(\zeta_P), [H_1(\hat{\zeta}_P) + H_2(\zeta_P)] \psi_0(\hat{\zeta}_P) \Phi_0(\zeta_P)) \\ &\quad + \sum_{P'} (\varepsilon_{P'} \psi_0(\hat{\zeta}_{P'}) \Phi_0(\zeta_{P'}) \phi_{P'}, H_{12} \varepsilon_{P'} \psi_0(\hat{\zeta}_{P'}) \Phi_0(\zeta_P) \phi_P), \end{aligned} \quad (\text{A} \cdot 7)$$

in which we have used the symmetry property with respect to P . If we take into account the explicit form of H_{12} and introduce the Majorana exchange operator P_{st} ,

$$\begin{aligned} (\Psi_0, H\Psi_0) &= (\psi_0(\hat{\zeta}_P) \Phi_0(\zeta_P), [H_1(\hat{\zeta}_P) + H_2(\zeta_P) \\ &\quad + \sum V_0(\tau_{st}) (1 - \alpha + \alpha P_{st})] \psi_0(\hat{\zeta}_P) \Phi_0(\zeta_P)) . \end{aligned} \quad (\text{A} \cdot 8)$$

The equations (4·4a) and (4·4b) are derived from the relation $\delta(\Psi_0, H\Psi_0) = 0$ corresponding to (A·6a) and (A·6b).

References

- 1) M. Goldhaber and E. Teller, Phys. Rev. **74** (1948), 1046.
- 2) H. Steinwedel, J. H. D. Jensen and P. Jensen, Z. f. Nat. **5a** (1950), 413.
J. M. Araújo, Nuovo Cim. **XII** (1954), 780.
P. Morrison, Report to Photonuclear Conference (1953).
I. Bloch and Y. Hsieh, Phys. Rev. **96** (1954), 382.
- 3) J. L. Burkhardt, Phys. Rev. **91** (1953), 420 (L).
A. Reifman, Z. f. Nat. **8a** (1953), 505.
J. S. Levinger and D. C. Kent, Phys. Rev. **95** (1954), 418.
J. S. Levinger, Phys. Rev. **97** (1955), 122.
D. H. Wilkinson, Proceedings of the 1954 Glasgow Conference (1954) 161.
- 4) M. Ferentz, M. Gell-Mann and D. Pines, Phys. Rev. **92** (1953), 836 (L).
- 5) S. Fujii and S. Takagi, Prog. Theor. Phys. **14** (1955), 402; 405, (L).
U. L. Businaro and S. Gallone, Nuovo Cim. **1**, **6** (1955), 1285.
- 6) S. Tomonaga, Prog. Theor. Phys. **13** (1955), 467; 482.
- 7) K. L. Brown and R. Wilson, Phys. Rev. **93** (1954), 443.
- 8) S. Tomonaga's lecture at Research Institute for Fundamental Physics on May, 1955.
- 9) A. J. F. Siegert, Phys. Rev. **52** (1937), 787.
- 10) L. L. Foldy, Phys. Rev. **92** (1953), 178.
J. S. Levinger, An. Rev. of Nucl. Sci., Vol. 2 (1953), 105.
- 11) M. Soga, S. Iijima and M. Nogami, private communication.
- 12) T. Tamura and T. Miyazima, Prog. Theo. Phys. **15** (1956), 255
- 13) A. M. Lane, R. G. Thomas and E. P. Wigner, Phys. Rev. **98** (1955), 693.
- 14) J. M. Blatt and J. D. Jackson, Phys. Rev. **76** (1949), 18.
E. E. Salpeter, Phys. Rev. **82** (1951), 60.
- 15) R. Nathans and J. Halpern, Phys. Rev. **93** (1954), 437.
R. Montalbetti, L. Katz and J. Goldemberg, Phys. Rev. **91** (1953), 659.
- 16) J. S. Levinger and H. A. Bethe, Phys. Rev. **78** (1950), 115.
- 17) J. Goldemberg and J. L. Lopes, Nuovo Cim. **12** (1954), 817 (L).
- 18) A. Bohr and B. R. Mottelson, Dan. Mat. Fys. Medd. **27** (1953), No. 16.
- 19) V. de Sabbata and A. Sugie, Nuovo Cim. **X** (1956), 16.
- 20) K. Okamoto, Prog. Theor. Phys. **15** (1956), 75 (L).
- 21) D. C. Peaslee, Phys. Rev. **88** (1952), 812.
- 22) N. P. Heydenturg and G. M. Temmer, Phys. Rev. **100** (1955), 150.

Energy Loss of a Charged Particle Traversing Superconductors

Satio HAYAKAWA* and Kazuo KITAO**

*Research Institute for Fundamental Physics, Kyoto University, Kyoto

**Department of Physics, Osaka City University, Osaka

(Received April 16, 1956)

The energy loss of a charged particle traversing superconductors is calculated by taking account of the London equations of the superconductivity. A small amount of excess energy loss, which is of the order of 10^{-2} or less, is expected, in comparison with the energy loss in normal materials. As this is just the limit of observability, an accurate measurement of the energy loss may give us information on the value of the characteristic constant of a superconductor, λ , introduced by London. The excess loss relative to the normal energy loss is estimated as about $2m/c^2 n \lambda$, where m is the mass of an electron and n the density of electrons.

§ 1. Introduction

In their recent paper Ivanenko and Tsytovich¹⁾ have discussed the energy loss of charged particles traversing a ferromagnetic material. The contribution of the magnetic permeability to the energy loss was found to be negligibly small, i.e., only of the order of 10^{-8} . This can be expected, because the magnetic permeability is appreciable only at very low frequencies of the electromagnetic field caused by a charged particle, whereas the main contribution to the energy loss comes from its optical frequencies. The present authors²⁾ have also considered similar problems, independent of Ivanenko and Tsytovich, paying particular attention to the superconductor. As the influence on the energy loss is much larger in the latter case than in the case of ferromagnetic materials, we want to publish our result, hoping that observations of the energy loss, if accurate enough, may serve to determine a constant characteristic to the superconducting state.

The energy loss of a charged particle in superconductors is different from that in normal matter in the following respects. It is well known that the energy loss due to electronic excitation is proportional to n/m , where n is the density of electrons and m the mass of an electron. In a superconductor the so-called superelectron behaves in a different way from normal electrons, so that the effective mass of the former, m_s , is different from m . Hence the contribution from superelectrons is considered to be proportional to n_s/m_s , where n_s is the density of superelectrons. As it is believed to be $m_s \lesssim 10^{-3} m$, one might think that superelectrons would contribute to the energy loss great deal, even though n_s is far smaller than the density of whole electrons, n . However, the characteristic behaviour of superelectrons almost disappears in the optical region where the main contribution to the energy loss comes from normal electrons; superelectrons play an essential role at lower

frequencies. This means that the characteristic frequency for superelectrons, $\omega_s = \sqrt{4\pi e^2 n_s / m_e}$, is smaller than that for normal electrons, $\omega_n = \sqrt{4\pi e^2 n / m}$. As we know that ω_s is smaller at least by one order than ω_n , the contribution of superelectrons relative to that of normal electrons is estimated as $(\omega_s / \omega_n)^2$, that is of the order of 10^{-2} or less. This is an effect which can barely be observed.

In addition to the effect discussed above, the perfect diamagnetism in superconductor shields the electromagnetic field produced by a charged particle. This may modify the density effect, but its magnitude seems to be negligibly small, because the modification takes place in the argument of a logarithm that appears in the formula of the energy loss.

The same conclusion can be obtained by the other way based on the London equations for the superconductivity. According to them superelectrons decrease the real part of the dielectric constant, so that the energy loss increases by the order of $(\omega_s / \omega_n)^2$. If conduction electrons play a special role in a microwave region, as suggested by Landau³⁾ their contribution to the energy loss is practically neglected, because they are essentially inert for the absorption of light at optical frequencies.

In § 2 the formula of the energy loss is derived by a conventional method employed, for example, by Budini⁴⁾. The formula is expressed, so that the characteristic constant of a superconductor, $\Lambda = 4\pi / \omega_c^2$, introduced by London appears explicitly. In § 3 numerical results are discussed for some typical superconductors. For Sn the energy loss at zero temperature is estimated as larger by 1.4% than that in the normal state, provided that the special role of conduction electrons is not operative. As this is just at the limit of observability nowadays, our case is not purely academic, in contrast to the case of Ivanenko and Tsytoich.

In Appendix is given the formula of the energy loss in magnetic materials. Although this is essentially the same as that derived by Ivanenko and Tsytoich, our formula applies to the case where both the dielectric constant and the magnetic permeability are different from unity.

§ 2. Derivation of formula

Let us consider a particle of charge $Z_1 e$ travelling along the z axis with velocity $c\beta$. Its charge and current densities at time t are expressed as

$$\rho^0 = Z_1 e \delta(x_0) \delta(y_0) \delta(z_0 - c\beta t), \quad \mathbf{j}^0 = c\beta \rho^0, \quad (2.1)$$

where x_0 , y_0 and z_0 indicate the position of the particle at $t=0$. The Fourier components of ρ^0 and \mathbf{j}^0 are given of angular frequency ω by

$$\rho_\omega^0 (Z_1 e / 2\pi c\beta) \delta(x_0) \delta(y_0) \exp(i\omega z_0 / c\beta), \quad \mathbf{j}_\omega^0 = c\beta \rho_\omega^0. \quad (2.2)$$

Our aim is to obtain the electromagnetic field induced by the charge and current given above.

In our case where the medium is a superconductor, the electromagnetic field is governed by the London equations. Namely, the magnetic and electric field strengths, \mathbf{H} and \mathbf{E} , are related to the supercurrent density \mathbf{j}_s as

$$c \operatorname{rot} (A \mathbf{j}_s) = -\mathbf{H}, \quad A \dot{\mathbf{j}}_s = \mathbf{E}, \quad (2.3)$$

where A is the characteristic constant of a superconductor and is assumed to be independent of position and frequency. The Maxwell equations together with (2.3) give us

$$\nabla^2 \phi_\omega + \epsilon_\omega (\omega^2/c^2) \phi_\omega = - (4\pi/\epsilon_\omega) \rho_\omega^0, \quad (2.4a)$$

$$\nabla^2 A_\omega + \epsilon_\omega (\omega^2/c^2) A_\omega = - (4\pi/c) \mathbf{j}_\omega^0, \quad (2.4b)$$

where the magnetic permeability is assumed to be unity. ϕ_ω and A_ω are the Fourier components of the scalar and electromagnetic potentials respectively. ϵ_ω here is different from the dielectric constant due to normal electrons, ϵ_ω^n , by

$$\epsilon_\omega = \epsilon_\omega^n - 4\pi/A\omega^2. \quad (2.5)$$

The solutions of (2.4a, b) are easily obtained in the cylindrical coordinate (z, r, θ) as

$$\phi_\omega = (Z_1 e / \pi c \beta \epsilon_\omega) K_0(k_\omega r) (i\omega z / c \beta), \quad (2.6a)$$

$$A_\omega = (Z_1 e \beta / \pi c \beta) K_0(k_\omega r) \exp(i\omega z / c \beta), \quad (2.6b)$$

where K_i is the modified Bessel function of the second kind of order i . The wave number k_ω is defined as

$$\begin{aligned} k_\omega^2 &= (\omega^2/c^2 \beta^2) (1 - \beta^2 \epsilon_\omega^n) + 4\pi/c^2 A \\ &= (\omega^2/c^2 \beta^2) (1 - \beta^2 \epsilon_\omega). \end{aligned} \quad (2.7)$$

From (2.6) yield the non-vanishing components of the field strengths

$$E_{z\omega} = - \frac{i Z_1 e \omega}{\pi c^2 \beta^2 \epsilon_\omega} (1 - \beta^2 \epsilon_\omega) K_0(k_\omega r) e^{i\omega z / c \beta} = \frac{i Z_1 e \omega}{\pi c^2} \frac{k_\omega^2}{k_\omega^2 - \omega^2/c^2 \beta^2} K_0(k_\omega r) e^{i\omega z / c \beta}, \quad (2.8a)$$

$$E_{r\omega} = \frac{Z_1 e k_\omega}{\pi c \beta \epsilon_\omega} K_1(k_\omega r) e^{i\omega z / c \beta} = - \frac{Z_1 e \omega^2}{\pi c^3 \beta} \frac{k_\omega}{k_\omega^2 - \omega^2/c^2 \beta^2} K_1(k_\omega r) e^{i\omega z / c \beta}, \quad (2.8b)$$

$$H_{\theta\omega} = \frac{Z_1 e k_\omega}{\pi c} K_1(k_\omega r) e^{i\omega z / c \beta}. \quad (2.8)$$

The electromagnetic field thus produced forms a Poynting vector, consequently an energy flow, that is equal to the energy lost by the charged particle. The energy flow passing through the surface of a cylinder of radius r is calculated as

$$\begin{aligned} - \frac{dW}{dz_0} \Big|_{>r} &= - \frac{r}{c \beta} \int_{-\infty}^{\infty} E_z H_\theta dz \\ &= \left(\frac{Z_1 e}{c \beta} \right)^2 \frac{r}{\pi} \int_0^\infty i \omega d\omega \left[\frac{k_\omega^*}{\epsilon_\omega} (1 - \beta^2 \epsilon_\omega) K_0(k_\omega r) K_1(k_\omega^* r) - \text{comp. conj.} \right]. \end{aligned} \quad (2.9)$$

Usually r is chosen so small that $|k_\omega| r \ll 1$ and, therefore, K_0 and K_1 are expressed by

their asymptotic forms

$$K_0(k_w r) \simeq \ln(2/\gamma k_w r), \quad K_1(k_w r) \simeq 1/k_w r, \quad (2 \cdot 10)$$

where $\ln \gamma$ is the Euler constant. Then (2.9) is reduced to

$$-\frac{dW}{dz_0}\Big|_{>r} = \frac{2}{\pi} \left(\frac{Z_1 e}{c\beta} \right)^2 \int_0^\infty \omega d\omega \left[\frac{\epsilon_{2w}}{|\epsilon_w|^2} \ln \left(\frac{2}{\gamma |k_w| r} \right) + \frac{1}{2} \left(\frac{\epsilon_{1w}}{|\epsilon_w|^2} - \beta^2 \right) \varphi \right], \quad (2 \cdot 10)$$

where ϵ_{1w} and ϵ_{2w} are the real and imaginary parts of ϵ_w ; namely,

$$\epsilon_w = \epsilon_{1w} + i\epsilon_{2w} = \epsilon_{1w}^n - 4\pi/\Lambda\omega^2 + i\epsilon_{2w}^n, \quad (2 \cdot 11)$$

φ is the phase angle of k_w^2 , as defined by

$$\tan \varphi \equiv \Im_m k_w^2 / \Re k_w^2 = -\beta^2 \epsilon_{2w} / (1 - \beta^2 \epsilon_{1w}) = -\beta^2 \epsilon_{2w}^n / (1 - \beta^2 \epsilon_{1w}^n + 4\pi\beta^2/\Lambda\omega^2). \quad (2 \cdot 12)$$

Usually $\tan \varphi$ is close to zero, so that φ can be approximated, according to two branches of \tan^{-1} , as

$$\varphi = -\beta^2 \epsilon_{2w} / (1 - \beta^2 \epsilon_{1w}), \quad \text{for } \beta^2 \epsilon_{1w} < 1, \quad (2 \cdot 13B)$$

$$\varphi = \beta^2 \epsilon_{2w} / (1 - \beta^2 \epsilon_{1w}) - \pi, \quad \text{for } \beta^2 \epsilon_{1w} > 1. \quad (2 \cdot 13C)$$

The two branches correspond to the Bohr and Cerenkov frequencies respectively. According as the frequency belongs to the Bohr or the Cerenkov one, the mechanism of the energy loss is caused by the excitation or the Cerenkov radiation. If the two mechanisms are written separately, (2.10) is expressed approximately as

$$\begin{aligned} -\frac{dW}{dz_0}\Big|_{>r} &= \frac{2}{\pi} \left(\frac{Z_1 e}{c\beta} \right)^2 \int_B \omega d\omega \frac{\epsilon_{2w}}{|\epsilon_w|^2} \left[\ln \left(\frac{2}{\gamma |k_w| r} \right) - \frac{\beta^2}{2} \epsilon_{1w} \right] \\ &\quad + \left(\frac{Z_1 e}{c\beta} \right)^2 \int_C \left(\beta^2 - \frac{\epsilon_{1w}}{|\epsilon_w|^2} \right) \omega d\omega. \end{aligned} \quad (2 \cdot 14)$$

The suffices B and C indicate that the integrals should be taken over the Bohr and Cerenkov frequencies respectively.

The derivation of (2.14) from (2.6) is exactly the same as that in the ordinary case⁴⁾. This part is included for the sake of completeness and also for the convenience of the readers. The important point in our case is implied entirely in the material constant ϵ_w , in which the characteristic constant Λ is contained. Unfortunately, however, the correct expression of ϵ_w is yet unknown. It is hoped that this will be obtained by experiments of the energy loss. Their possibility is discussed on the following assumptions.

First we adopt the assumption proposed by Landau³⁾. According to Landau, the conduction electrons of density n_e turn out below the critical temperature, T_c , to be superconductors in part and the rest of them are frozen in certain bound states, because otherwise the normal conduction electrons would contribute to the finite electric resistivity. As the electrons begin to freeze at temperature T_c , their characteristic frequency is regarded as $\omega_c \sim kT_c/\hbar$. Whole electrons, whose density is Zn_0 , are thus classified into the following

three kinds; the bound electrons of density $Zn_0 - n_c$ form such oscillators that have frequencies ω_j , strengths f_j and damping frequencies g_j ; the frozen electrons of density $n_c - n_s$ are represented by an oscillator of frequency ω_c and of damping frequency g and their effective mass, m_f , which may practically be equal to the electronic mass, m ; the superelectrons of density n_s and of effective mass m_e take part in ϵ_ω as in (2.11), these being related to Λ as

$$\omega_s^2 \equiv 4\pi/\Lambda = 4\pi e^2 n_s / m_e. \quad (2.15)$$

With these quantities the dielectric constant is expressed as

$$\epsilon_\omega = 1 + \frac{4\pi e^2 (Zn_0 - n_c)}{m} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - ig_j \omega} + \frac{4\pi e^2 (n_c - n_s)}{m_f (\omega_c^2 - \omega^2 - ig \omega)} - \frac{\omega_s^2}{\omega^2}. \quad (2.16)$$

It may be useful to give the orders of magnitude of the frequencies in (2.16).

$$\omega_j = 10^{16} \sim 10^{19} \text{ sec}^{-1}, \quad \omega_c = 10^{11} \sim 10^{12} \text{ sec}^{-1}, \quad \omega_s \sim 10^{16} \text{ sec}^{-1}. \quad (2.17)$$

Needless to say, ω_c , ω_s and n_s decrease as temperature approaches T_c and vanish above T_c .

If Landau's postulate were rejected, we would have to amalgamate the third term in the right hand side of (2.16) into the second one, so that

$$\epsilon_\omega = 1 + \frac{4\pi e^2 (Zn_0 - n_s)}{m} \sum_j \frac{f_j}{\omega_j^2 - \omega^2 - ig_j \omega} - \frac{\omega_s^2}{\omega^2}, \quad (2.16')$$

where the summation over j extends to conduction electrons. Comparing this with (2.16) one can see that (2.16') is a special case of (2.16).

Now we consider the excitation loss, the Bohr part, of (2.14). The integral over ω is evaluated at resonance frequencies of $\epsilon_\omega/|\epsilon_\omega|^2$. On account of the numerical values given in (2.17), the contribution from a resonance due to the frozen electrons is by far, by a factor of about 10^{-5} , smaller than that from the bound electrons. Therefore, we have only to take account of the optical resonances; the main contribution comes from a frequency region of $\omega \sim \omega_j$. The density of superelectrons, n_s , can be neglected in comparison with n_c , on account of $n_s \lesssim 10^{-3} n_c$. Similarly n_s in (2.16') can be neglected in comparison with Zn_0 . Putting $m_f = m$, therefore, the second and the third terms in the right hand side of (2.16) can be put together like as the second term in the right hand side of (2.16'), because the presence of ω_c is practically ineffective for the energy loss. Thus the postulate proposed by Landau will hardly be examined, unless the accuracy of observing the energy loss becomes extremely good.

The above considerations lead us to conclude that the superelectrons will be only a source of modifying the energy loss to detectable extent. The formula of the excitation loss is then given as

$$\begin{aligned} -\frac{dW}{dz_0} \Big|_{< r, B} &= \frac{2\pi Z_1^2 Zn_0 e^4}{mc^2 \beta^2} \sum_j \frac{f_j}{|1 - \omega_s^2 / \bar{\omega}_j^2| \epsilon_{\omega_j}^2} \\ &\times \left[\ln \frac{4c^2 \beta^2}{\gamma^2 |1 - \beta^2 \epsilon_{\omega_j}^2 \bar{\omega}_j^2| r^2} - \beta^2 \epsilon_{\omega_j}^2 \left(1 - \frac{\omega_s^2}{\bar{\omega}_j^2 \epsilon_{\omega_j}^2} \right) \right], \end{aligned} \quad (2.18)$$

where $\bar{\omega}_j$ is the actual resonance frequency of the j -th level and is only slightly different from ω_j . The effect of the superconductivity comes in $\omega_s^2/\bar{\omega}_j^2$. Such a term in the square bracket of (2.18) is negligible, because the logarithmic term plays a dominant role. The modification of $\epsilon_{\bar{\omega}_j}$ in the logarithm is also negligible. Therefore, the modification due to the superconductivity lies in the denominator $|1 - \omega_s^2/\bar{\omega}_j^2 \epsilon_{\bar{\omega}_j}^n|$, that gives a positive contribution to the energy loss.

In evaluating the additional part, we may safely put $\epsilon_{\bar{\omega}_j}^n = 1$ and take $\bar{\omega}_j^2$ out of \sum_j by replacing it by an average value

$$\langle \bar{\omega}_j^2 \rangle = 4\pi e^2 Z n_0 / m. \quad (2.19)$$

Then the ratio of the additional excitation loss in a superconductor to the excitation loss in a normal material can be expressed as

$$\Delta \equiv 2\omega_s^2 / \langle \bar{\omega}_j^2 \rangle = 2n_s m / Z n_0 m_e. \quad (2.20)$$

This is what we expected on the basis of a qualitative consideration in § 1.

Next we consider the Cerenkov radiation represented by the last term of (2.14). The influence of the term $\omega_s^2/\bar{\omega}_j^2$ increases $\epsilon_{\omega}/|\epsilon_{\omega}|^2$ but decreases ϵ_{ω} both by $\omega_s^2/\bar{\omega}_j^2$. Hence both the intensity and the critical velocity for the Cerenkov radiation increase slightly, though such effects may hardly be detectable in an optical region. However, an important contribution to the Cerenkov radiation may arise from a microwave region, provided that Landau's postulate holds. According to Landau, the dielectric constant at such frequencies is extremely large, $10^7 \sim 10^{10}$, so that even a particle of non-relativistic velocity can radiate the electromagnetic wave. However, the intensity of such a radiation is very weak and its contribution to the energy loss is negligibly small.

§ 3. Numerical results and discussions

Before giving the numerical values of Δ in (2.20) we have to notice that ω_s or Δ is temperature dependent. Let us introduce a quantity that refers to zero temperature:

$$P = n_{s0} m / n_0 m_e, \quad (3.1)$$

where n_{s0} is the density of superelectrons at zero temperature. Adopting the current assumption on the temperature dependence of n_s ,

$$n_s = n_{s0} [1 - (T/T_c)^4], \quad (3.2)$$

for temperature T below the critical temperature T_c , the temperature dependence of the fractional excess excitation loss Δ is expressed as

$$\Delta = (2/Z) P [1 - (T/T_c)^4]. \quad (3.3)$$

Keeping the temperature dependence (3.3) in mind, we shall discuss the magnitude of Δ at zero temperature, Δ_0 . The quantity P is obtained by observing the penetration depth of electromagnetic field in superconductors⁵⁾. Based on such data, the values of Δ_0 for four superconductors are shown in Table 1.

Table 1. Fractional excitation loss at zero temperature

| material | In | Sn | Hg | Pb |
|----------------|------|------|------|------|
| Z | 49 | 50 | 80 | 82 |
| P | 0.22 | 0.35 | 0.30 | 0.70 |
| $\Delta_0(\%)$ | 0.90 | 1.4 | 0.75 | 1.7 |

The numerical values of Δ_0 given in Table 1 tell us that the fractional excess loss is of the order of percent or less, which is barely larger than a detectable limit for Sn. However, it may not be easy to increase the accuracy by another order, so that the accuracy of P or Δ obtained from the energy loss seems a little worse than that from the penetration depth. Nevertheless, the former mean is regarded as superior over the latter, because of such a reason that the sample suitable for the observation of the penetration depth is difficult to prepare and the measurement might be sensitive to the sample, while the observation of the energy loss is practically independent of the preparation of a sample. On the other hand, the energy loss measurement suffers from such a difficulty that the energy lost by charged particles turns to heat up the sample so as to destroy the superconductivity. This requires to decrease the beam intensity of charged particles as low as possible, but not so low as to be masked by the background. With regard to this situation, we think it valuable to undertake the observation of the energy loss in superconductors, in order to see their physical properties.

Appendix.

Energy loss in magnetic materials

If both the dielectric constant and the magnetic permeability are different from unity, we obtain the formula for the energy loss as follows. The magnetic permeability that depends on frequency ω is expressed, analogously to ϵ_ω in (2.11), as

$$\mu_\omega = \mu_{1\omega} + i\mu_{2\omega}. \quad (A.1)$$

Then the energy loss due to the region of a cylinder of radius r is

$$\begin{aligned}
 -\frac{dW}{dz_0}\bigg|_{>r} &= \frac{2}{\pi} \left(\frac{Z_1 e}{c\beta} \right)^2 \int_B \omega d\omega \frac{\epsilon_{2\omega}}{|\epsilon_\omega|^2} \left[\ln \frac{2}{\gamma |k_\omega| r} - \frac{\beta^2}{2} \epsilon_{1\omega} \mu_{1\omega} \right] \\
 &\quad + \frac{2}{\pi} \left(\frac{Z_1 e}{c\beta} \right)^2 \beta^2 \int_B \omega d\omega \mu_{2\omega} \left[\ln \frac{2}{\gamma |k_\omega| r} - \frac{1}{2} \right] \\
 &\quad + \left(\frac{Z_1 e}{c\beta} \right)^2 \int_C \omega d\omega \left(\beta^2 \mu_{1\omega} - \frac{\epsilon_{1\omega}}{|\epsilon_\omega|^2} \right), \quad (A.2)
 \end{aligned}$$

where

$$k_\omega^2 = (\omega^2/c^2\beta^2) (1 - \beta^2 \epsilon_\omega \mu_\omega). \quad (A.3)$$

The suffices B and C indicate that the integrals over ω extend over the Bohr and Cerenkov

frequencies respectively ; namely,

$$\beta^2 \mathcal{R}e(\epsilon_\omega \mu_\omega) < 1 \quad (B), \quad \beta^2 \mathcal{R}e(\epsilon_\omega \mu_\omega) > 1 \quad (C).$$

The numerical result given by (A, 2) is essentially the same as that obtained by Ivanenko and Tsytovich.

References

- 1) D. Ivanenko and V. N. Tsytovich, J. Exp. Theor. Phys. USSR **28** (1955), 291.
- 2) S. Hayakawa and K. Kitao, Lecture at the Physical Society of Japan (1954).
- 3) Cited in D. Shoenberg, Nuovo Cimento, Supplemento **10** (1953), 458.
- 4) P. Budini, Nuovo Cimento, **10** (1953), 3.
- 5) D. Shoenberg, Superconductivity, Camb. Univ. Press, 2nd ed. (1952).

Energy Loss of a Charged Particle Traversing Ionized Gas and Injection Energies of Cosmic Rays

Satio HAYAKAWA* and Kazuo KITAO**

* *Research Institute for Fundamental Physics, Kyoto University, Kyoto*

***Department of Physics, Osaka City University, Osaka*

(Received April 16, 1956)

The energy loss of a charged particle traversing an electron gas is calculated by taking the direct collisions with free electrons and the plasma excitation into account. The contributions from ions and neutral atoms are added and the resultant energy loss is given as a function of the degree of ionization. The energy loss is found to increase with the degree of ionization, to which excess loss is approximately proportional. The critical injection energy of cosmic rays, based on the Fermi mechanism of acceleration, is given in Table 1. Its value is about fifty percent larger for the degree of ionization of ten percent than for a neutral medium and nearly four times larger for a half ionized medium. In the case of the solar production of cosmic rays the critical injection energy is below one MeV per nucleon, so that the secondary particles generated by nuclear reactions could be accelerated to cosmic ray energies. The mechanism of acceleration in a coronal region is proposed in the Appendix to explain the time delay of the cosmic ray increase compared with the solar eruption as well as the energy spectrum of solar cosmic rays.

§ 1. Introduction

In some astrophysical problems, it is important to know the energy loss of a charged particle in a highly ionized gas. This fact has been emphasized in a number of occasions, for example, by Morrison, Olbert and Rossi¹⁾ in connection with the injection of cosmic rays and by Greenstein²⁾ in discussing the role of electrons in the stellar atmosphere. These authors had to use the formula of the ionization loss for a neutral medium or its approximate modification, because no reliable knowledge of that for an ionized medium was available. Recently, however, Neufeld and Ritchie³⁾ have worked out this problem by paying attention to the excitation of plasma waves. Independently of them the present authors⁴⁾ have also treated the same problem in a simple way. In this paper we present our method of derivation as well as its numerical results which may serve to astrophysical problems, and discuss the bearing of our formula in the injection of cosmic rays.⁵⁾

In deriving the formula of the energy loss in an electron gas, we use the method of impact parameter. For small impact parameters the energy loss is regarded as due to the direct collision with electrons, while for large impact parameters as due to the excitation of plasma waves. The two regions are divided by a critical impact parameter which is of the order of the Debye length. Adding the contributions from the respective regions, the critical impact parameter is eliminated from the final formula.

The energy loss due to free electrons thus derived is added to the contribution from ions and neutral atoms. Then the total energy loss is expressed in terms of the degree of ionization, α . The fractional excess energy loss is found to be of the order of α/Z , where Z is the atomic number of the medium concerned, and to be nearly independent of energy in the relativistic region, while it increases slowly with decreasing energy in the non-relativistic region.

In order that our result is applicable to astrophysical problems, numerical values are presented for the cases of the interstellar space and of the solar corona, in which gases are mostly composed of hydrogen with the degrees of ionization of ten percent and of hundred percent, respectively. In the case of the interstellar gas, we also show the dependence of the energy loss upon the degree of ionization, expecting the application to the galactic halo, in which the degree of ionization is supposed to be higher. The effect of ionization of a gas is found considerable for the solar corona. The energy loss of a proton is so large that it can not penetrate the corona, if it would start from the bottom of the corona with the velocity of auroral particles. The large ionization loss sets the lower limit of injection energies of cosmic rays from the sun as below 1 MeV, provided that the Fermi mechanism of acceleration is operative.

§ 2. Derivation of formula

Let the density of atoms with atomic number Z be n_0 . For the degree of ionization, α , the density of neutral atoms is $(1-\alpha)n_0$ and their contribution to the energy loss is given by

$$-dW/dz_0|_n = 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot Z(1-\alpha)n_0 \\ \times \sum_j f_j \left[\ln \frac{4m^2 c^4 \beta^4 \Gamma^2}{r^2 |1 - \beta^2 \epsilon_{\omega_j}| \hbar^2 \bar{\omega}_j^2} - \beta^2 \epsilon_{1\omega_j} \right], \quad (2.1)$$

where $\Gamma = (1 - \beta^2)^{-1/2}$ and other notations are the same as used in our preceding paper⁹⁾. There are ions of density αn_0 which are assumed to be singly ionized. The extension to the case of higher degree of ionization is straightforward. Then the contribution of those ions is

$$-dW/dz_0|_i = 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot (Z-1)\alpha n_0 \\ \times \sum_j f_j \left[\ln \frac{4m^2 c^4 \beta^4 \Gamma^2}{r^2 |1 - \beta^2 \epsilon_{\omega_j}| \hbar^2 \bar{\omega}_j^2} - \beta^2 \epsilon_{1\omega_j} \right]. \quad (2.2)$$

The summation over j in (2.2) is a little different from that in (2.1), but the difference is insignificant for large Z .

The ionization produces free electrons of density αn_0 . Their contribution is what we want to calculate. For small values of impact parameter, $r < r_0$, where r_0 is of the order of the Debye length, the energy loss is considered as due to the direct collision with such free electrons. This gives rise to the energy loss:

$$-dW/dz_0|_{<r_0} = 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot \alpha n_0 \ln(m^2 c^2 \beta^2 r_0^2 / \hbar^2). \quad (2.3)$$

For large values of impact parameter, $r > r_0$, on the other hand, the energy loss is caused by the excitation of plasma waves. The energy loss can then be calculated, if the dielectric constant of the electron plasma is known. This is

$$\epsilon_\omega = 1 - \omega_p^2 / (\omega^2 - i\omega\nu), \quad (2.4)$$

where ν is the collision frequency of an electron in the medium and ω_p the plasma frequency given by

$$\omega_p^2 = 4\pi\alpha n_0 e^2 / m \equiv \alpha\omega_0^2. \quad (2.5)$$

With use of (2.4) we obtain the energy loss for the region concerned as

$$\begin{aligned} -dW/dz_0|_{>r_0}^e &= 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot \alpha n_0 \\ &\times \left[\ln \frac{4c^2 \beta^2}{\gamma^2 |1 - \beta^2 \epsilon_{\omega p}| r_0^2 \omega_p^2} - \beta^2 \epsilon_{1\omega p} \right]. \end{aligned} \quad (2.6)$$

Adding (2.3) and (2.6), the energy loss due to free electrons is obtained as

$$\begin{aligned} -dW/dz_0|^e &= 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot \alpha n_0 \\ &\times \left[\ln \frac{4m^2 c^4 \beta^4 \Gamma^2}{\gamma^2 |1 - \beta^2 \epsilon_{\omega p}| \hbar^2 \omega_p^2} - \beta^2 \epsilon_{1\omega p} \right]. \end{aligned} \quad (2.7)$$

It must be noticed that r_0 does not appear in the final formula (2.7).

(2.7) can be compared with formula (87) obtained by Neufeld and Ritchie³⁾, who used a more complicated method, as shown in what follows. In practical cases both the density of atoms is not too large and the degree of ionization is not too small, so that there holds

$$\omega_p^2 = \alpha\omega_0^2 \gg \nu^2.$$

Hence $\epsilon_{\omega p}$ and $\epsilon_{1\omega p}$ can be assumed as zero. Thus we have

$$-dW/dz_0|^e = 2\pi Z_1^2 e^4 / mc^2 \beta^2 \cdot \alpha n_0 \ln(4m^2 c^4 \beta^4 \Gamma^2 / \gamma^2 \hbar^2 \omega_p^2). \quad (2.7')$$

This is nothing but their formula (87), if one notices $b_{min} = \hbar / mc\beta\Gamma$ there.

Adopting (2.7') instead of (2.7) for practical purpose, the dependence on α is expressed as $-\alpha \ln \alpha$. In a similar way we write down the α dependence of the total energy loss, the sum of (2.1), (2.2), and (2.7'), as

$$-\frac{dW}{dz_0} = A \left[B + \frac{\alpha}{Z} (C - B - \ln \alpha) \right], \quad (2.8)$$

where

$$\begin{aligned} A &= 2\pi Z_1^2 e^4 / m c^2 \beta^2, \\ B &= \sum_j f_j \left[\ln \frac{4m^2 c^4 \beta^4 \Gamma^2}{\gamma^2 |1 - \beta^2 \epsilon_{1\bar{\omega}_j}| (\hbar \bar{\omega}_j)^2} - \beta^2 \epsilon_{1\bar{\omega}_j} \right], \end{aligned} \quad (2.9)$$

$$C = \ln \frac{4m^2 c^4 \beta^4 \Gamma^2}{\gamma^2 \hbar^2 \omega_0^2}. \quad (2.9)$$

(2.8) indicates that the maximum of the energy loss appears for $C-B-1 < 0$ at

$$\alpha_{max} = \exp(C-B-1). \quad (2.10)$$

Introducing $\bar{\omega}$, the average of $\bar{\omega}_j$ over oscillators, we can approximate

$$C-B-1 \simeq \ln[|1 - \beta^2 \epsilon_{\bar{\omega}}| (\bar{\omega}^2 / \bar{\omega}_0^2)] - \epsilon_{\bar{\omega}} \beta^2 - 1. \quad (2.11)$$

As $\bar{\omega}^2 \gg \omega_0^2$ in usual cases, (2.11) is always positive, so that the energy loss increases with α in a monotonic way. The energy dependence of the excess energy loss in the ionized medium is due mainly to $|1 - \beta^2 \epsilon_{\bar{\omega}}|$ in (2.11). This tends to increase with decreasing energy; the difference of the energy loss between in ionized and neutral media is larger for lower energies. Since the logarithm in (2.11) is as large as B , the fractional excess energy loss is roughly estimated as α/Z .

§ 3. Numerical results for hydrogen gas

In most celestial cases the gas is mainly composed of hydrogen. Then (2.8) with $Z=1$ is of practical importance and is expressed as

$$-\frac{dW}{dz_0} = \frac{2\pi n_0 Z_1^2 e^4}{mc^2 \beta^2} \left[\left\{ \ln \frac{4m^2 c^4 \beta^4}{\gamma^2 (1 - \beta^2) [1 - \beta^2 \epsilon_{\bar{\omega}}] \hbar^2 \bar{\omega}^2} - \beta^2 \right\} + \alpha \left\{ \ln \frac{(1 - \beta^2) \bar{\omega}^2}{\omega_0^2 \alpha} + \beta^2 \right\} \right]. \quad (3.1)$$

Numerical examples of (3.1) are illustrated for the interstellar gas and the solar corona.

In the galactic plane the interstellar gas is composed mainly of hydrogen with the degree of ionization of about $\alpha=1/10$. The average density of hydrogen is supposed to be approximately $n_0=1/2 \text{ cm}^{-3}$. This results in $\omega_0 \sim 4 \times 10^4 \text{ sec}^{-1}$, far smaller than $\bar{\omega} \simeq 2.1 \times 10^{16} \text{ sec}^{-1}$. Owing to the small density $\epsilon_{\bar{\omega}}$ in (3.1) can be set equal to unity.

On account of that α may change from place to place, we express (3.1) with α as a parameter. In a medium of density $\rho \text{ g cm}^{-3}$, the energy loss is given by

$$-\frac{1}{\rho} \frac{dW}{dz_0} = \frac{0.153 Z_1^2}{\beta^2} \left[\left\{ 21.3 + 2 \ln \frac{\beta^2}{1 - \beta^2} - \beta^2 \right\} + \alpha \{ \ln(1 - \beta^2) + \beta^2 - \ln \alpha + 53.9 \} \right] \text{ MeV/g cm}^{-2}. \quad (3.2)$$

For $Z_1=1$ and $\alpha=0.1$ and 0.5 , (3.2) is plotted against the kinetic energy of a proton in Fig. 1. For comparison the energy loss for $\alpha=0$ is also shown in the same figure. In comparison with the latter case the energy loss for $\alpha=0.1$ increases nearly by twenty percent. The excess loss is nearly independent of energy in a relativistic region, while it increases weakly with decreasing energy in a non-relativistic region. This can be understood, on account of that the plasma frequency is low, while the atomic frequency is so high that a part of the electric field caused by a charged particle becomes ineffective, as its velocity decreases.

In order to show the dependence on the degree of ionization, the α dependence at

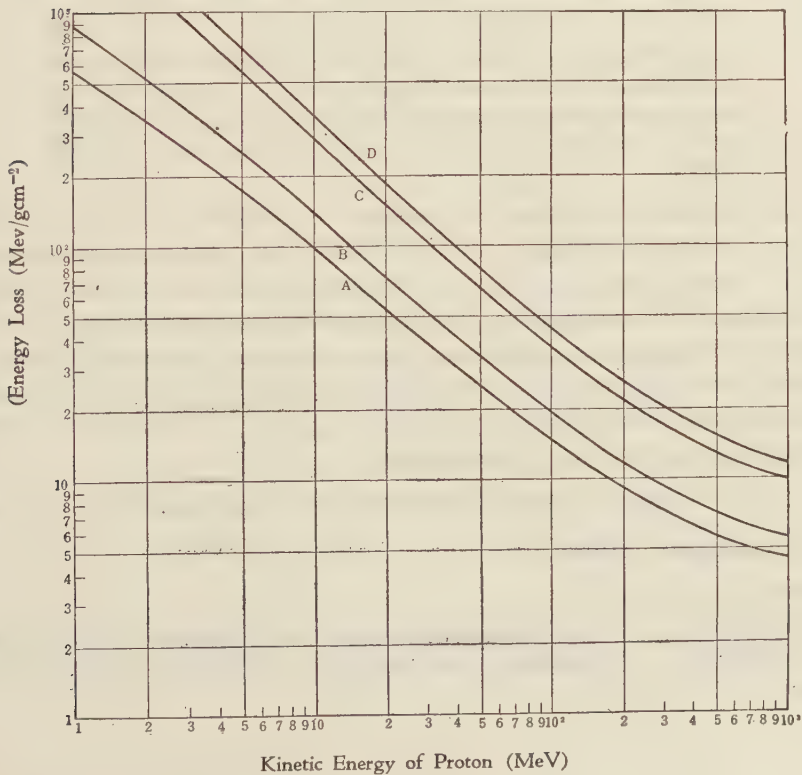
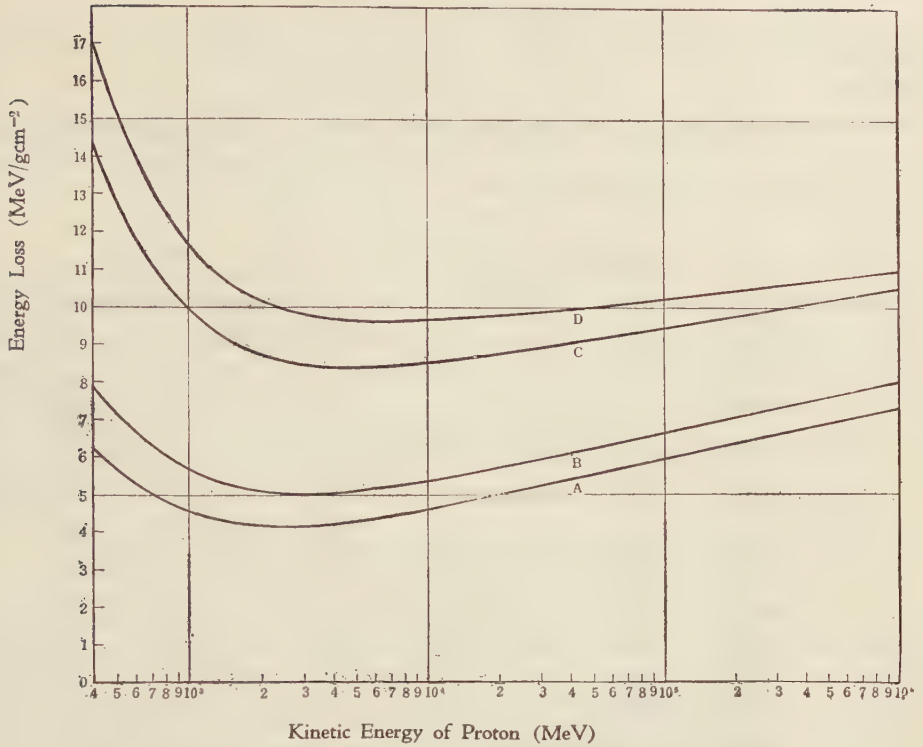


Fig. 1. Energy loss (MeV/g cm^{-2}) vs the kinetic energy of proton (MeV). Curves A, B and C correspond to the hydrogen gas of density $1/2 \text{ cm}^{-3}$ with $\alpha=0, 0.1$ and 0.5 , respectively. Curve D corresponds to the solar corona with $\alpha=1$.

$W=2.18 \text{ GeV}^{*)}$ for protons, at which the energy loss is minimum in a neutral medium, is shown as

$$-1/\rho \cdot dW/dz_0|_{min} = 0.168[25.0 + \alpha(52.4 - \ln \alpha)] \text{ MeV/g cm}^{-2}. \quad (3.3)$$

(3.3) is plotted in Fig. 2, together with its ratio to the case of $\alpha=0$.

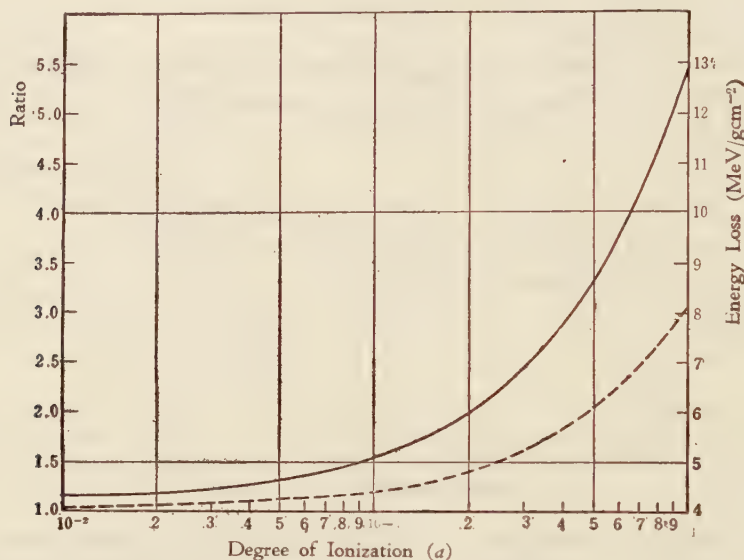


Fig. 2. Energy loss *vs* degree of ionization at $W=2.18 \text{ GeV}$ for proton in the hydrogen gas of density $1/2 \text{ cm}^{-3}$.

Solid curve indicates the energy loss (MeV/g cm^{-2} represented in the right ordinate) and dashed curve the ratio to the neutral medium (in the left ordinate).

In the solar corona the degree of ionization is supposed to be unity. Hence the energy loss is obtained by putting $\alpha=1$ in (3.1). The numerical result is shown also in Fig. 1. The density of electrons is of the order of 10^8 cm^{-3} which is still small enough to neglect the density effect. The electron temperature is so high as $10^6 \text{ }^\circ\text{K}$ that the motion of electrons plays a significant role in the energy loss of slow particles. Since the average velocity of electrons in the corona is about $3 \times 10^8 \text{ cm sec}^{-1}$, however, the effect of the electron motion may be neglected for protons of energies above 1 MeV. In a MeV region the energy loss is approximated as

$$-1/\rho \cdot dW/dz_0 = 144/W \cdot [\ln W + 22.3] \text{ MeV/g cm}^{-2} \quad (3.4)$$

where W is measured in MeV.

*) By W we mean the kinetic energy, while E is used for the total energy. W and E with the same suffix, as employed in § 4 and Appendix, are related as, say, $E_c = Mc^2 + W_c$, where Mc^2 is the rest energy of a proton.

Neglecting the electron motion which is of little practical effect at energies concerned, the range of protons can be estimated in the following way. We assume that (3.4) holds down to the energy at which the velocity of a proton is equal to the thermal velocity of electrons. The range for a proton with energy W_p in MeV to lose its energy down to this energy is approximately given by

$$R(W_p) \simeq W_p^2 / 5530 \text{ g cm}^{-2}. \quad (3.5)$$

The range for a proton below this energy is negligible for practical purpose. As the depth of the corona is of the order of $10^{-5} \text{ g cm}^{-2}$, the proton of energy below 200 KeV can not penetrate the corona. If the spiral motion of protons in a magnetic field is taken into account, the minimum energy of penetration will be larger. The protons of currently believed velocity, about 10^8 cm sec^{-1} , as the cause of magnetic storms, can not at all penetrate the corona. If such particles were injected from the sun, they would have to be accelerated near the top of the corona.

§ 4. Injection energy of cosmic rays

According to the Fermi mechanism of acceleration of cosmic rays, cosmic ray particles are injected into the interstellar space, in which wandering magnetic clouds take part in the acceleration of the particles. In order that the acceleration is effective, the particles have to be injected with energies larger than a critical injection energy, the energy at which the gain and loss of energy of a particle are equal. This is given by Morrison et al.¹⁾ on the assumption that the cosmic radiation is stored in our galaxy of disk shape. This assumption seems, however, to be in contradiction with recent evidences in cosmic rays as well as in astronomy. As described by one of the present authors (S. H.)²⁾, it seems better to regard the stored region as spherical, containing the galactic halo, in which the average thickness of matter for cosmic rays to traverse is of the order of $x \simeq 1 \text{ g cm}^{-2}$ *). Then the critical injection energy is, as in (3.5) of reference 7, given by

$$E_c = 1.5 \beta(W_c) x, \quad (4.1)$$

where factor 1.5 comes from the exponent of the power energy spectrum of cosmic rays. $\beta(W_c)$ is the energy loss per g cm^{-2} for a particle of energy W_c . The value of W_c has been computed by employing the energy loss in a neutral medium¹⁾⁶⁾⁷⁾. Now we are able to get a more realistic estimate of W_c , based on our calculation.

The critical injection energies per nucleon are given in Table 1 for $\alpha=0.1$ and 0.5 and for some typical cosmic ray nuclei, assuming $x=1 \text{ g cm}^{-2}$. These values of W_c are, of course, larger than in the case of the neutral medium. The ionization loss for $\alpha=0.5$ is appreciable, so that the energy spectrum at low energy would be affected considerably,

*) The thickness may be different for protons, because a proton has a higher momentum than heavy nuclei for the same rigidity. On account of that the mechanism of escape from the galaxy is not solved yet, however, we use the same value of x for all nuclei.

if the mean path length traversed were larger by one order.

Table 1. Critical injection energy (W_c/A in MeV)

| α | Z_1 | 1 | 2 | 4 | 7 | 14 | 26 |
|----------|-------|------|------|-----|-----|----|----|
| | A | 1 | 4 | 9 | 14 | 28 | 56 |
| 0 | | 0.95 | 0.95 | 2.2 | 4.8 | 11 | 20 |
| 0.1 | | 1.6 | 1.6 | 3.5 | 7.2 | 16 | 29 |
| 0.5 | | 4.5 | 4.5 | 8.5 | 16 | 32 | 65 |

Actually the average energy of cosmic ray particles injected into the interstellar space is supposed to be much higher than the critical injection energy W_c . As discussed in our previous paper⁷⁾, cosmic rays suffer acceleration in the vicinity of local sources, such as Crab nebula. Also from the sun particles of energies above 1 GeV are generated. Such high energies may also be explained in terms of the Fermi mechanism that is operative near local sources, where physical conditions are significantly different from those in the general interstellar space. For example, the density of matter, the degree of ionization, the velocity of magnetic clouds and the strength of magnetic fields may be much higher in the vicinity of a local source than in the general interstellar space. Hence the acceleration in the former region is considered as more efficient than in the latter. And it should also be subjected to the critical injection energy that will intimately be related to an instantaneous process of acceleration, such as the betatron mechanism. Thus we presume that the acceleration of cosmic rays can be divided into three steps, the instantaneous injection from local sources, efficient acceleration near the local sources and the gradual acceleration in the general interstellar space. As we have been concerned with the last step in the preceding paragraphs, the second step will be discussed in the Appendix.

Based on (A.5) with (A.6) the critical injection energy in the solar corona is approximately given by

$$E_1 = 3 \times 10^{13} \beta(W_1) \rho. \quad (4.2)$$

Here ρ is the density of matter in the corona and is supposed to be of the order of $10^{-16} \text{ g cm}^{-3}$. As the corona is completely ionized, α is taken as unity in (3.1). The values of W_1 per nucleon are found to be below 1 MeV, at which our formula may hardly be applicable, because of the influence of rapid electron motion.

It seems to us worth remarking that the critical injection energy for protons is so low that a certain fraction of protons from $d-d$ reactions may contribute to cosmic rays. These reactions possibly supply a sufficient amount of energy to the disturbance occurring on the solar surface, as suggested by Greenstein.²⁾ Furthermore, this might cause such a stream of high velocity, $\sim 10^{10} \text{ cm sec}^{-1}$, that could cause the solar radio emission of Type III observed by Wild et al.⁸⁾ In this connection it will be very interesting to study the intercorrelation between the energy supply and the particle acceleration which will take place in

the stellar atmosphere.

Appendix

An acceleration mechanism of solar cosmic rays

Let the dimension of a region be R , in which the acceleration in the second step is operative. Inside this region there are numerous magnetic clouds which move with the average velocity V . The average strength of magnetic fields in the clouds is represented by H . Owing to the magnetic field the particles injected undergo the random walk with transport mean free path l , which is as large as the average dimension of the magnetic clouds. H and l determine a rigidity, above which the rigidity spectrum becomes steep. As we are mainly concerned with protons and electrons, the rigidity is equal to the momentum divided by unit charge. Thus in the unit of eV/c the corresponding momentum is given by

$$P_1 \simeq 300 Hl/c, \quad (\text{A} \cdot 1)$$

where H and l are measured in gauss and cm respectively. The particles with momenta below P_1 suffer collisions with the magnetic clouds before they escape out of the region. The number of collisions is roughly estimated as

$$N \simeq (R/l)^2. \quad (\text{A} \cdot 2)$$

Hence the average lifetime of particles for escape is

$$t \simeq Nl/v \simeq R^2/lv, \quad (\text{A} \cdot 3)$$

where v is the velocity of particles and is nearly equal to the light velocity c . A collision with a magnetic cloud results in the rate of energy gain by $2(V/v)^2$, so that the exponent of the power energy spectrum turns out to be

$$\gamma \simeq 1/N2(V/v)^2 \simeq (l/R)^2/2(V/v)^2. \quad (\text{A} \cdot 4)$$

In the medium of density ρ the average thickness of traversed matter is $Nl\rho \simeq tv\rho$. It is almost always satisfied that this quantity is much smaller than the mean free path for nuclear collisions. The critical injection energy, W_1 , in this case is given by

$$E_1 = \gamma\beta(W_1) Nl\rho \simeq \gamma\beta(W_1) tv\rho, \quad (\text{A} \cdot 5)$$

analogous to (4.1).

Now we apply the above considerations to the solar production of cosmic rays. Among various quantities introduced above, P_1 , t and γ are directly related to observations. The energy spectrum of the solar cosmic rays becomes steep above several GeV. Below this the energy spectrum is not so steep and γ is estimated as nearly unity. The lifetime for escape may be estimated on the time interval between the solar eruption and the cosmic ray increase as about ten minutes.⁹⁾ Therefore we take the numerical values of P_1 , t and γ as

$$P_1 = 10^{10} \text{ eV}/c, \quad t = 10^3 \text{ sec}, \quad \gamma = 1. \quad (\text{A} \cdot 6)$$

These values indicate only the order of magnitude and are subject to ambiguities of factor two or so. In addition to these values, we assume R as large as the radius of the coronal region, namely

$$R=10^{11} \text{ cm.} \quad (\text{A} \cdot 7)$$

Four quantities given in (A·6) and (A·7) are sufficient to determine others, if v is assumed as equal to c . Thus we have

$$l \simeq 3 \times 10^8 \text{ cm,}^{*)} \quad H \simeq 0.1 \text{ gauss,} \quad V \simeq 7 \times 10^7 \text{ cm sec}^{-1}. \quad (\text{A} \cdot 8)$$

These values seem to be reasonable in view of the following facts: l is as large as the thickness of a shock front generated by a stream of velocity V ; the magnetic field strength on the solar surface is about one gauss and the equipartition between magnetic and kinetic energies predicts H of a fraction of gauss; the velocity of particle streams associated with solar disturbances is believed as about 10^8 cm sec^{-1} . However, there arises a question on the behaviour of electrons. The magnetic field strength given in (A·8) allows to accelerate the electrons by the Fermi mechanism up to about several GeV, above which the energy loss due to the synchrotron radiation forbids the acceleration. The spectrum of radio waves generated by these electrons may have a maximum intensity at frequency of about 10^{13} cyc/sec which is larger than the observed frequencies of solar disturbed radio waves, around 10^8 cyc/sec . However, the mechanism of the solar radio burst is due likely to the plasma oscillation taking place in the corona⁽⁶⁾⁽¹⁰⁾ and the synchrotron radiation may be relatively unimportant as its cause.

References

- 1) P. Morrison, S. Olbert and B. Rossi, *Phys. Rev.* **94** (1954), 191.
- 2) J. L. Greenstein, Conference on Radio Astronomy at Manchester (1955); The authors are grateful to Professor Hatanaka who let them know the report of Greenstein
- 3) J. Neufeld and R. H. Ritchie, *Phys. Rev.* **98** (1955), 1632.
- 4) S. Hayakawa and K. Kitao, Lecture at the Physical Society of Japan (1954).
- 5) Our fundamental idea is based on E. Fermi, *Phys. Rev.* **75** (1949), 1169.
- 6) S. Hayakawa and K. Kitao, *Prog. Theor. Phys.* **16** (1956), 131.
- 7) S. Hayakawa, *Prog. Theor. Phys.* **15** (1956), 111.
- 8) J. P. Wild, J. D. Murray and W. C. Rowe, *Aust. J. Phys.* **7** (1954), 439.
- 9) H. Elliot, *Progress in Cosmic Ray Physics* **1** (1952), 455.
- 10) S. Hayakawa and N. Hokkyo, *Prog. Theor. Phys.* **15** (1956), 193.

*) The value of l is regarded as fixed. So the P_1 per nucleon for heavy nuclei is a half of that for protons.

Letters to the Editor

An Attempt of Generalizing the Hypothesis of Charge Independence

Osamu Hara and Yasunori Fujii

*Institute of Theoretical Physics,
Nagoya University, Nagoya*

May 28, 1956

The hypothesis of charge independence has played an important role in the phenomena of pion and of heavy unstable particles. Its validity is, however, restricted, and one can not apply it to decay processes. In this letter is proposed an attempt which, by generalizing the concept of the isotopic spin, aims to evade this restriction.

We introduce angular momenta \mathbf{J} and \mathbf{K} in charge space other than the usual isotopic spin \mathbf{I} ,* and assume that the total isotopic spin is given by

$$\mathbf{L} = \begin{cases} \mathbf{I} + \mathbf{J} + \mathbf{K} & \text{for fermion,**} \\ \mathbf{I} + \mathbf{J} & \text{for boson.} \end{cases}$$

The charge is assumed to be given by eL_3 . Therefore J_3 and K_3 are related to the

strangeness S and the particle number n introduced by Gell-Mann and Nishijima¹⁾ by

$$J_3 = S/2,$$

and

$$K_3 = n/2.$$

Here we assume

- (i) L_3 is conserved in all interactions,
- (ii) in interactions where electromagnetic field does not take part, $\mathbf{I} + \mathbf{J}$ is conserved,
- (iii) in strong interactions \mathbf{I} and \mathbf{J} are separately conserved.

(iii) corresponds to the assumption of usual charge independence. But it is in general more stringent in that it requires the conservation of \mathbf{J} .

(i) is nothing but the conservation law of charge, but (ii) and (iii) give results which have not been got heretofore. We can predict, for example, the branching ratios of some decay processes or the relations between lifetimes of some unstable particles. This is because, as long as we assume that the decay interactions are primary interactions and the electromagnetic field does not take part there, $\mathbf{I} + \mathbf{J}$ is conserved according to (ii). The calculation is elementary, and the main result is as follows :*

* One reason for this generalization is that the gauge transformation can be regarded as a rotation in charge space around its third axis. If this is taken as serious, we must assume that the conservation law of charge must have the form of the conservation law of the third component of an angular momentum in charge space.

** As for fermion we consider baryons exclusively at first.

* Recently Kawaguchi and Nishijima, Iso, Takeda and Wentzel²⁾ obtained the same result under the assumption that for decay processes a selection rule $\Delta I = \frac{1}{2}$ holds. The agreement of our result with theirs is obvious since in the processes here considered $\Delta J = \frac{1}{2}$ and the conservation of $\mathbf{I} + \mathbf{J}$ just leads to $\Delta I = \frac{1}{2}$.

- (i) Branching ratio for the decay of Λ^0 : $\Lambda^0 \rightarrow p + \pi^- / \Lambda^0 \rightarrow n + \pi^0 = 2$,
 (ii) Lifetime of Σ : $\tau_{\Sigma^+} \leq 3\tau_{\Sigma^-}$,
 (iii) Lifetime of θ : $\tau_{\theta^\pm} = \frac{1}{2}\tau_{\theta^0}$
 if θ is (odd spin, odd parity),

and the decay mode

$$\left. \begin{array}{l} \theta^\pm \rightarrow \pi^\pm + \pi^0 \\ \text{is forbidden, and the branching} \\ \text{ratio for the decay of } \theta^0: \\ \theta^0 \rightarrow \pi^+ + \pi^- / \theta^0 \rightarrow 2\pi^0 = 2 \end{array} \right\}$$

if θ is (even spin, even parity).

(iii) seems to reject the possibility (odd spin, odd parity), since recent experiments show definitely that τ_{θ^\pm} is about one hundred times longer than τ_{θ^0} .⁽³⁾ If this is the case, the decay of θ^\pm must be ascribed to interactions entirely different from those which are responsible for the decay of θ^0 .

These predictions can be compared with experiments directly, and the validity of our assumptions may be checked by comparing them with experiments.

Details will be published in a later issue of this journal.

- 1) M. Gell-Mann, Phys. Rev. **92** (1953), 833.
K. Nishijima, Prog. Theor. Phys. **12** (1954), 107.
- 2) M. Kawaguchi and K. Nishijima, Soryushiron kenkyu (mimeographed circular in Japanese) **9** (1955), 449. Prog. Theor. Phys. **15** (1956), 180, 182.
T. Iso, *ibid.* **11** (1956), 1.
G. Takeda, Phys. Rev. **101** (1956), 1547.
G. Wentzel, Phys. Rev. **101** (1956), 1215.
- 3) For example, B. Rossi, Proceedings of the Fifth Rochester Conference (1955).
As for recent data on K^+ , see L. W. Alvarez, F. S. Crawford, M. L. Good and M. L. Stevenson, Phys. Rev. **101** (1956), 503.

A Conjecture on the Possible Existence of New Hyperons

Osamu Hara and Yasunori Fujii

*Institute of Theoretical Physics,
Nagoya University, Nagoya*

May 28, 1956

As assumed in the preceding letter, I and J are conserved in strong interactions in our scheme. Therefore the hamiltonian of the system must commute with these observables (except for electromagnetic and weak interactions). The form of the hamiltonian is restricted by this condition, and it is seen that it can contain I and J only in the form I^2 and J^2 . If we assume that it has the form

$$H = \alpha \cdot p + \beta M(I^2, J^2),$$

the eigenvalue of $M(I^2, J^2)$ corresponds to the rest mass. Assuming most simply that it is a linear combination of I^2 and J^2 , and determining their coefficients from the masses of the nucleon, Λ and Ξ , it is given by (in electron mass unit)

$$M \sim 100 I(I+1) + 400 J(J+1) + 1800 \quad (1)$$

This formula reproduces the hyperon masses within the error of $\sim 100 m_e$.

(1) gives the mass separation of two levels with I differing by one and the same J , and that with J differing by one-half and the smallest possible I , as

$$\left. \begin{array}{l} \Delta_I M = 200 I, \\ \Delta_J M = 400 J + 100 + (-1)^J 75. \end{array} \right\} \quad (2)$$

If

$$\Delta_I M \leq m_\pi \quad \text{and} \quad \Delta_J M \leq m_\pi, \quad (3)$$

Table 1. Level scheme for heavy hyperons

| I | J | mas. | charge | | | | | |
|---------------|---------------|-------------|----------------------|--|--|---|---|--------------------|
| | | | $-2e$ | $-e$ | 0 | e | $2e$ | $3e$ |
| $\frac{1}{2}$ | 2 | ~ 4300 | $(-\frac{1}{2}, -2)$ | $(-\frac{1}{2}, -1)$ $(\frac{1}{2}, -2)$ Z^- | $(\frac{1}{2}, -1)$ $(-\frac{1}{2}, 0)$ | $(-\frac{1}{2}, 1)$ $(\frac{1}{2}, 0)$ | $(\frac{1}{2}, 1)$ $(-\frac{1}{2}, 2)$ Z^{++} | $(\frac{1}{2}, 2)$ |
| 1 | $\frac{3}{2}$ | ~ 3500 | $(-1, -\frac{3}{2})$ | $(-1, -\frac{1}{2})$ $(0, -\frac{3}{2})$ | $(-1, \frac{1}{2})$ $(0, -\frac{1}{2})$ $(1, -\frac{3}{2})$ Y^0 | $(0, \frac{1}{2})$ $(1, -\frac{1}{2})$ $(-1, \frac{3}{2})$ Y^+ | $(1, \frac{1}{2})$ $(0, \frac{3}{2})$ | $(1, \frac{3}{2})$ |
| 0 | $\frac{3}{2}$ | ~ 3500 | | $(0, -\frac{3}{2})$ X^- | $(0, -\frac{1}{2})$ | $(0, \frac{1}{2})$ | $(0, \frac{3}{2})$ | |

The levels denoted by the dotted lines are unstable. Figures in the parentheses indicate I_3 and J_3 respectively.

Table 2. Production and decay of heavy hyperons

| hyperons | production scheme | corresponding threshold energy (Bev) | possible decay scheme* |
|-----------------|---------------------------------|--------------------------------------|------------------------------------|
| Z^{++}, Z^- | $\pi + N \rightarrow Z + 4K$ | 8.9 | $Z \rightarrow X(Y) + \pi$ |
| | $N + N \rightarrow Z + Z$ | 8.3 | $\rightarrow \Xi + K$ |
| | $\rightarrow Z + N + 4K$ | 11.9 | $\rightarrow \Lambda(\Sigma) + 2K$ |
| | etc. | | etc. |
| Y^+, Y^0, X^- | $\pi + N \rightarrow Y(X) + 3K$ | 4.7 | $Y(X) \rightarrow \Xi + \pi$ |
| | $N + N \rightarrow Y(X) + Y(X)$ | 4.2 | $\rightarrow \Lambda(\Sigma) + K$ |
| | $\rightarrow Y(X) + N + 3K$ | 7.3 | etc. |
| | etc. | | |

* The selection rule $\Delta J = \frac{1}{2}$ was used to deduce this decay scheme.

the level can be metastable. (3) can be written, using (2), in the form

$$I \leq 1 \text{ and } J \leq 2. \quad (4)$$

This means that the number of hyperons other than those which have been established heretofore will not be so large. Those $I \leq 1$ and $J \leq 1$ are already known (nucleon, Λ , Σ and Ξ).^{*} Some features of higher levels will be discussed below.

a) $J = 3/2$ and $I = 0$ and 1.

(1) gives $M \simeq 3300$ and 3500 for $I = 0$ and 1 respectively. At $I = 0$ four, and at $I = 1$ twelve levels appear (see the

* A number of unwanted levels appear in this scheme (for example, Σ^{++}). Most of them, however, can be eliminated by discussions similar to those made in this letter, and we arrive at the result which is almost the same as the ones proposed by Gell-Mann,¹⁾ Pais¹⁾ and Nishijima.²⁾ As for the detail see our full paper which will be published soon.

table). It can be shown, however, that most of them decay rapidly into lower levels by emitting γ -rays or pions, and the remaining levels are X^- , Y^0 and Y^+ only. Their threshold energies and possible decay schemes are listed in the table. As will be seen, it seems possible to identify the new hyperon reported by Eisenberg³⁾ as our X^- .

b) $J=2$ and $I=1/2$.

(1) gives $M \simeq 4300$. Eliminating unstable levels as before, we get Z^{++} and Z^- . Recent observations by Fry et al.⁴⁾ that K^- emerges from a star can be interpreted in terms of Z^- .^{*} But our mass formula is not so accurate, and Z lies near the border of instability. Therefore, we can not say positively whether such particles really exist.

- 1) M. Gell-Mann, Phys. Rev. **92** (1953), 833.
M. Gell-Mann and A. Pais, Proc. Glasgow Conf. (1954).
- 2) K. Nishijima, Prog. Theor. Phys. **12** (1954), 107.
- 3) Y. Eisenberg, Phys. Rev. **96** (1954), 641. Proc. Pisa Conf. (1955).
- 4) W. F. Fry and M. S. Swami, Phys. Rev. **97** (1955), 1189.
W. F. Fry, J. Schneps and M. S. Swami, Proc. Pisa Conf. (1955).

* X , Y and Z are unstable in the nucleus, since following rapid processes

$$X(Y) + N \rightarrow \Xi + A$$

$$Z + N \rightarrow 2\Xi, \text{ etc.}$$

occur. Therefore Fry's event must, if interpreted in terms of ours, be the capture of X^- or Z^- . Strong interaction of X leading to the emission of K is energetically impossible, and the only possibility is the capture of Z . The capture of Z^- can lead to the emission of K^- according to

$$Z^- + N \rightarrow \Xi + A + K^-.$$

It is a remarkable feature of this picture that Ξ and A must be accompanied to K .

Inelastic Scattering of Nucleons by Nuclei

Ryoichi Kajikawa, Tatuya Sasakawa*
and Wataro Watari

Department of Physics, Hiroshima University,
Hiroshima and

Department of Physics, Kyoto University,*
Kyoto

May 21, 1956

Austern, Butler and McManus¹⁾ put forward a theory of inelastic and/or exchange scattering. They used the plane waves for the wave functions of incoming and outgoing nucleons and they evaluated the matrix element for the scattering due to the contribution only from outside the nucleus. Such a treatment may give good results when the energy of the incident particle is so high that the use of the Born approximation is justified.

However, it has been shown at energies lower than, say, 15 Mev²⁾ that a considerable part of the interaction between incident or outgoing nucleons and nuclear matter can be described as a form of an average potential.³⁾ This suggests that the distortion of the nucleon wave inside the nucleus is not negligible. The imaginary part appearing in the average potential (i.e. optical potential) gives a measure of nuclear excitation. But it has been suggested that this is relatively small compared with the real part of the optical potential. Therefore it is expected that a perturbation method may be applicable to a calculation of the inelastic scattering of a nucleon by the nucleus. And the gross structure of

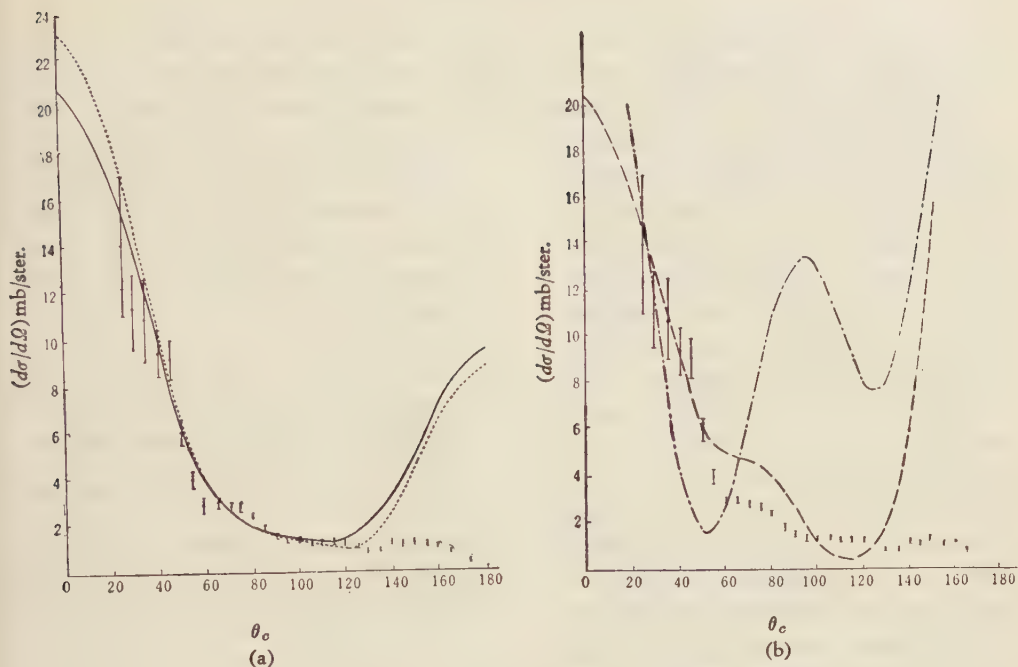


Fig. 1. Comparison of theoretical and experimental results

| form of $V(i\ n)$ | depth of the average potential | configuration of the final state of Fe^{56} |
|--------------------|--------------------------------|--|
| — : delta function | 30 Mev | $(f_{7/2})^6(2)$ |
| - - - : Yukawa | 30 Mev | |
| — : delta | 45 Mev | $(f_{7/2})^5(7/2)p_{3/2}(3/2)(2)$ |
| - - - : delta | 45 Mev | |

Configuration of the ground state of Fe^{56} is assumed to be $(f_{7/2})^6(0)$. Theoretical curves are normalized at 50° for (a) and at 30° for (b) to the experimental value. It can be shown that the spin exchange force and the tensor force gives no contributions in $0^+ \rightarrow 2^+$ transition in first order perturbation calculation

the angular distribution of the inelastic process would be reproduced by a calculation taking into account the distortion of incoming and outgoing waves due to the optical potential, and treating the interaction giving rise to excitation itself as a perturbation.

As an example we take a case of protons inelastically scattered by Fe^{56} leaving it in its first excited state. Calculations are made for an incident energy of 17 Mev and the results are compared with the experiment of Schrank et al.⁴⁾ To get

the scattering amplitude, we calculated a matrix element,

$$\langle \Phi^{(-)} \chi_\varepsilon | V | \chi_{\varepsilon_0} \Psi^{(+)} \rangle,$$

where $\Psi^{(+)}$ is the outgoing wave solution of the Schrödinger equation, and $\Phi^{(-)}$ is the ingoing wave solution of the particle subject to the optical potential.⁵⁾ χ_{ε_0} and χ_ε are wave functions of ground and excited states of the target nucleus of Fe^{56} , respectively. V is the sum of two body interactions

$$V = \sum_{n=1}^A V(i\ n).$$

Here i indicates the incident particle. We adopt partial wave analysis taking account of partial waves up to six. This is mainly because the Racah coefficients applicable to the partial waves of orders higher than six are not available in conventional tables. If we take higher waves into account we shall get better results than what indicated below, since the momentum transfer becomes smaller for those higher waves which are neglected in the present calculation. We approximate $\Psi^{(+)}$ by an elastically scattered wave which is distorted by the same optical potential as in the case of $\Phi^{(-)}$. By the time, we take a real square well of depth 30 Mev or 45 Mev and of range $1.45A^{1/3} \times 10^{-13}$ cm, as the optical potential. The integration extends over the whole space.

The results are shown in Fig. 1. All these curves are normalized at 50° . These figures show that

- I) the main feature of the angular distribution is reproduced by our calculation,
- II) the angular distribution is strongly dependent on an assumed configuration of the nucleus. After having an adequate knowledge on the optical potential, we could determine the configuration to some extent from the angular distribution.
- III) We get a result nearly equal to the one with the point interaction and the finite range interaction for $V(i n)$.
- IV) A large forward scattering appears in all of our calculated distributions. The reason why Austern et. al.¹⁾ could not explain this forward peak seems to be due to their neglect of distortion of the proton waves.
- V) The backward rise in our curve will be noticed. One reason for this discrepancy will be the neglect of partial waves higher than six in the present calculation. The

other reason may be the neglect of the effect due to particle exchange.

Detailed account of the analysis will be published shortly in this journal.

We wish to express our sincere thanks to the Professor Sakuma of Hiroshima University, Professor Kobayasi of Kyoto University and their colleagues for their hospitality.

- 1) Austern, Butler and McManus, *Phys. Rev.* **92** (1953), 350.
- 2) For example, LeLevier and Saxon, *Phys. Rev.* **87** (1952), 40.
- 3) Feshbach, Porter and Weisskopf, *Phys. Rev.* **92** (1953), 448.
- 4) Schrank, Gugelot and Dayton, *Phys. Rev.* **96** (1954), 1156.
- 5) S. Hayakawa and S. Yoshida, *Prog. Theor. Phys.* **14** (1954), 1.

On the Theory of Thermal Conductivity of Pure Metals at Low Temperatures

P. G. Klemens

Division of Physics, Commonwealth Scientific and Industrial Research Organization, Sydney, Australia

June 18, 1956

The thermal conductivity of metals is proportional to

$$(\epsilon, c) \equiv \int_{-\infty}^{+\infty} \epsilon \frac{1}{(\epsilon^2 + 1)(1 + e^{-\epsilon})} c(\epsilon) d\epsilon \quad (1)$$

where $c(\epsilon)$ is the solution of the Bloch equation, which, in the limit of low temperatures and in the absence of scatter-

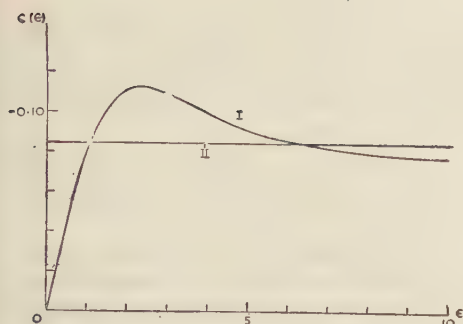


Fig. 1. Solution of equation (2). Curve I: Numerical solution according to Klemens. Curve II: Kasuya's trial function (5), normalized ($\alpha=0.85$).

ing by lattice imperfections, is of the form

$$S. c \equiv \int_0^{\infty} \frac{x^2 dx}{e^x - 1} \left\{ [c(\varepsilon + x) - c(\varepsilon)] \frac{e^{\varepsilon} + 1}{e^{\varepsilon} + e^{-x}} + [c(\varepsilon - x) - c(\varepsilon)] \frac{e^{-\varepsilon} + 1}{e^{-\varepsilon} + e^{-x}} \right\} = \varepsilon, \quad (2)$$

-see, for example, Sommerfeld and Bethe.¹⁾ It is easily seen that $c(\varepsilon)$ must be an odd function of ε .

It is also known^{2),3)} that

$$(S. c, c) \equiv \int_{-\infty}^{+\infty} S. c(\varepsilon) \frac{1}{(e^{\varepsilon} + 1)(1 + e^{-\varepsilon})} \times c(\varepsilon) d\varepsilon \quad (3)$$

is a maximum for variations in $c(\varepsilon)$, subject to the normalizing condition

$$(S. c, c) = (\varepsilon, c). \quad (4)$$

Kohler²⁾ obtained a value of (1), using the trial function $c = a\varepsilon$, and adjusting a by (4); Sondheimer⁴⁾ effected a considerable improvement by using $c = a\varepsilon + b\varepsilon^3$, and adjusting the parameters by maximizing (3), subject to (4).

Since it is easily seen that the solution of (2) cannot be represented by the first few terms of a power series in ε ,

Klemens⁵⁾ solved (2) directly by numerical methods. His solution is shown in Figure 1. By a variety of tests it was concluded that this solution is a good representation of the true functional form of $c(\varepsilon)$, and that any error in the resulting value of (ε, c) is well below $\frac{1}{2}\%$. The value so derived is 11% greater than Sondheimer's.

Kasuya⁶⁾ also noted that the solution of (2) could not be expressed by a power series, and suggested as a trial function

$$\begin{aligned} c(\varepsilon) &= -\alpha & \varepsilon > 0 \\ &= +\alpha & \varepsilon < 0, \end{aligned} \quad (5)$$

Adjusting α by (4) he obtained a value $\alpha = 0.105$.

The resulting value of the conductivity exceeds that derived from Klemens' numerical solution by almost 13%. This is anomalous if the estimated accuracy of that solution is accepted, for according to the variational principle the better trial function should yield the higher value of (ε, c) .

The author has therefore repeated the calculation of α and obtained a value of $\alpha = 0.0849$. The resulting value of (ε, c) is now 9% lower than the value derived from the numerical solution, thus removing the anomaly. This new value of α has been subsequently confirmed by Kasuya (private communication).

It is interesting to note that, with this new value of α , (ε, c) is about 2% higher than Sondheimer's value. The adjusted trial function (5) is also shown in Figure 1, and can be compared with the numerical solution. It is seen that (5) is indeed a fair representation of $c(\varepsilon)$ over a wide range of ε , and could easily be superior to a trial function of the form $c(\varepsilon) = a\varepsilon + b\varepsilon^3$.

Since the trial function (5) has the advantage of very easy numerical manipulation, one would be inclined to use it at intermediate temperatures and in the presence of residual resistance, in order to test Sondheimer's solution for these cases. Unfortunately, as the temperature and residual resistance are increased, the true solution deviates progressively from the form (5), tending towards the form $c = a\varepsilon$, -see, for example, Klemens.⁷⁾ Thus the trial function (5) ceases to hold an advantage over Sondheimer's trial function in these cases. The case of lowest temperatures and no residual resistance, which has been discussed here, is the case most unfavourable to Sondheimer's method; even

then his value of (ε, c) differs only by 11% from the numerical value. In other cases his results must be even closer to the correct value.

The author wishes to thank Miss J. Ward for performing the computations.

- 1) A. Sommerfeld and H. Bethe, *Hanb. d. Physik*, vol. 24/2. (Berlin: Springer, 1933).
- 2) M. Kohler, *Z. f. Physik* **124**, (1948), 772; **125** (1949), 679.
- 3) A. H. Wilson, *Theory of Metals*, 2nd ed. (Cambridge: University Press, 1953).
- 4) E. H. Sondheimer, *Proc. Roy. Soc. A* **203** (1950), 75.
- 5) P. G. Klemens, *Aust. J. Phys.* **7** (1954), 64.
- 6) T. Kasuya, *Prog. Theor. Phys.* **13** (1955), 561.
- 7) P. G. Klemens, *Encycl. of Physics*, vol. **14** (Berlin: Springer, 1956).

On the Quantum Field Theory of the Interaction between the Graviton and the Matter Field

Toshiei KIMURA

*Research Institute for Theoretical Physics, Hiroshima University
(Takehara-machi, Hiroshima-ken)*

(Received April 30, 1956)

In order to quantize the gravitational field so as to be consistent with the general relativistic treatment, we investigate what physical effects the quantizable and unquantizable parts of gravitational field have, when the matter fields exist. To construct the Hamiltonian, we show that the interaction Hamiltonian in the interaction representation can be obtained by the method extended from that of Yang-Feldman. The role of the non-linear part of gravitational field for such a problem is also discussed. Only the potential, obtained by eliminating the longitudinal and scalar parts of gravitational field with the aid of the supplementary condition, has an effect upon the macroscopical phenomena and gives the equation of two-body problem in general relativity found by Einstein, Infeld and Hoffmann. On the other hand, while the pure transverse graviton does not contribute to the macroscopical phenomena, it gives, in the region of high energy, the large damping effect upon the interaction between the elementary particles.

§ 1. Introduction and summary

Recently, in connection with the difficulties encountered in the quantum field theory, the effect of the gravitational field upon the interaction between the elementary particles has been discussed by some authors. In order to discuss the gravitational field in the theory of quantized field, we must first consider what sort of the classical gravitational equation should be quantized, and secondly examine which part of the gravitational field quantities can be quantized and what role another part that cannot be quantized plays in the macroscopical phenomena.

In spite of many attempts on the quantization of the gravitational field, there are neither necessary reasons for the quantization nor experimental evidences for the existence of the gravitational quanta (say, graviton). It will not be reasonable, therefore, to discuss only the quantum character of gravitational field. In other words, the problem whether the field equations or field quantities that we quantize are appropriate, should be considered not merely by looking into the properties of graviton, but also by examining whether the procedure of quantization is compatible with the general relativistic treatment. For example, if we take Einstein's gravitational equation as the one that we quantize, the gravitational field must be quantized in such a way that the part remaining unquantized can give the same physical effects as in the general relativity.

In the linear approximation of Einstein's gravitational field equation, the metric

tensor satisfies the same type of equations as those of electromagnetic potentials. Accordingly, in the quantization procedure, there will be many analogies between the gravitational and electromagnetic fields. In the classical theory of electromagnetic field, the field quantities are composed of two parts: the transverse part for which energy and momentum form a four-vector and behave like the energy and momentum of particles as regards its transformation properties, and the part for which this is not the case. In the quantum electromagnetic field theory, only the first part is subject to a quantization, giving rise to photons which behave like particle; while the second remains unquantized. The latter part composed of longitudinal and scalar parts, in the presence of matters, gives rise to the Coulomb interaction between charged matters both in the classical and in the quantum electromagnetic theories.

Analogously, in the gravitational field it will be reasonable to quantize in such a way that the unquantized part gives the gravitational potential corresponding to the Coulomb potential in the quantum electrodynamics. Then, what potential should we take as the gravitational potential between matters? Since the unquantized part can be taken for both the macroscopical and microscopical phenomena, it will be reasonable to take a such one that contains not only Newton potential but also retarded potential and gives rise to the equation of two-body problem in general relativity which was found by Einstein, Infeld and Hoffmann¹⁾ (EIH equation). As is well known in the electrodynamics, the effect of retardation is caused by the part of interaction term which depends on the transverse electromagnetic waves; it appears as the mutual emission of light waves between the particles. But in the gravitational theory, as we are ignorant of the transverse gravitational wave even from a macroscopical standpoint, it seems natural to deduce the retarded potential without using any transverse gravitational wave. The primary aim of this paper is to show that the quantization of the gravitational field is possible by separating the field quantities in such a way that the above mentioned potential is obtained. And the secondary aim is to make clear what physical effects the graviton, obtained by quantizing the transverse gravitational field, has in the microscopical world.

For the quantization of Einstein's gravitational field, we use the coordinate condition of De Donder²⁾:

$$\partial(\mathfrak{G}^{\mu\nu})/\partial x_\nu = 0,$$

$$(\mathfrak{G}^{\mu\nu} = \theta^{\mu\nu} (-\text{Det}(\theta_{\alpha\beta}))^{1/2} = \varepsilon^{\mu\nu} - \kappa \eta^{\mu\nu}, \quad \varepsilon^{11} = \varepsilon^{22} = \varepsilon^{33} = -\varepsilon^{44} = -1, \quad \varepsilon^{\mu\nu} = 0 (\mu \neq \nu)) \quad (1.1)^{*,**}$$

which corresponds to the Lorentz condition in the electromagnetic field. Adopting this condition, Papapetrou³⁾ succeeded in putting the Einstein's equation in the simpler form:

$$\varepsilon^{\alpha\beta} \partial^2 (\mathfrak{G}^{\mu\nu}) / \partial x^\alpha \partial x^\beta = \kappa^2 \theta^{\mu\nu} \quad (1.2)$$

where $\theta^{\mu\nu}$ is the energy momentum tensor of the whole system including the gravitational

* Belinfante has recently discussed the meaning of this condition. See, reference 6).

** $\hbar = c = 1$. $\kappa^2 = 16\pi G = 4.1 \times 10^{-48} \times 3 \times 10^{10} \times 10^{-27} \text{ cm}^2 \sim 12 \times 10^{-65} \text{ cm}^2$.

field itself. Taking the field equations for a 'free' gravitational field as

$$\varepsilon^{\alpha\beta} \partial^2 (\mathcal{G}^{\mu\nu} / \partial x^\alpha \partial x^\beta) = 0 \quad (1.3)$$

we can regard the entire interaction as caused by $\theta^{\mu\nu}$.

In order to discuss the quantum field theory of the interaction between the gravitational field and matter field (including the electromagnetic field), it will be convenient to develop the theory in the Minkowski space which can be regarded as the zero-th order approximation to the curved Riemannian space, after Gupta.^{(4), (5)} By expanding all the field quantities in powers of κ , it is possible to express them as expansions in the Minkowski space. In other words, Einstein's theory can also be regarded as a theory of gravitation in flat space with a Lagrangian density containing an infinite number of terms. After passing over from the Riemannian space to the Minkowski space, we shall split the gravitational field into the linear and non-linear parts. Then we shall treat the linear part as a 'free' gravitational field, while the non-linear part is regarded as the direct interaction between the gravitons. (§ 2)

In § 3, we shall consider whether the Hamiltonian formalism for the graviton interacting with the matter field can exist or not. For this purpose, following the method extended from that of Yang-Feldman, we investigate whether the interaction representation can be constructed or not. Even when we start with the first term of the interaction Lagrangian density, the interaction hamiltonian in the interaction representation does not close and is expressed as an infinite series. In consequence of this situation, there occur several complicated features. For instance, there appears infinite series of normal dependent terms in the supplementary condition and different interaction Hamiltonians are obtained according to the β -formalism or wave formalism for the meson field, despite we start with the same interaction Lagrangian density. We shall show that the non-linear part plays an important role to remove these unsatisfactory features.

In § 4, by eliminating the longitudinal and scalar parts of the gravitational field interacting with matter field with the aid of the supplementary condition, we obtain the gravitational potential between the matters. It is shown that EIH equation can be deduced by using this potential.

Recently, by taking into account the coordinate condition (1.1), Belinfante⁽⁶⁾ has suggested, within the scope of classical theory, that the space has holes where matters are located. Further he observed that there is a natural cut-off radius for the fields around such a hole, which will remove the divergence difficulties met with in the quantum field theory. We consider, in § 5, how this circumstance is to be taken into account in our formalism. The interactions between the gravitons and the matter fields are the so-called interactions of the second kind,⁽⁷⁾ except those between the scalar graviton and the trace of energy momentum tensor of the matter field. From the fact that the interaction of the second kind has large field reaction, it is shown that we can obtain the non-singular quantum electrodynamics and meson theory through replacing the free propagation function of matter fields by the one corrected by the graviton.

§2. Lagrangian density and field equations

As Gupta has shown,¹⁾ Einstein's gravitational theory can be regarded as a theory of gravitation in flat space with a Lagrangian density containing an infinite number of terms.

When there is no interaction between the gravitational field and other matter fields, we can take the following Lagrangian density as the zeroth order approximation

$$L_G = -\frac{1}{4}(\partial_\lambda \gamma_{\mu\nu} \cdot \partial_\lambda \gamma_{\mu\nu} - 2\partial_\mu \gamma_{\lambda\nu} \cdot \partial_\nu \gamma_{\lambda\mu} - \frac{1}{2}\partial_\lambda \gamma \cdot \partial_\lambda \gamma) \quad (2.1)$$

where $\partial_\lambda = \partial/\partial x_\lambda$, $\gamma = \gamma_{\mu\mu}$ and $\gamma_{\mu\nu}$ is symmetrized.* The second term of (2.1) can be expressed as

$$\frac{1}{2}\partial_\mu \gamma_{\lambda\nu} \cdot \partial_\nu \gamma_{\lambda\mu} = \frac{1}{2}\partial_\nu (\partial_\mu \gamma_{\lambda\nu} \cdot \gamma_{\lambda\mu}) - \frac{1}{2}\gamma_{\lambda\mu} \cdot \partial_\mu \partial_\nu \gamma_{\lambda\nu}.$$

If we take the coordinate condition $\partial_\nu \gamma_{\mu\nu} = 0$, we can drop the second term on the right-hand side of (2.1), and we have

$$L_G = -\frac{1}{4}(\partial_\lambda \gamma_{\mu\nu} \cdot \partial_\lambda \gamma_{\mu\nu} - \frac{1}{2}\partial_\lambda \gamma \cdot \partial_\lambda \gamma) \quad (2.1')$$

Treating $\gamma_{\mu\nu}$ and γ as independent variables, (2.1') gives the energy momentum tensor $t_{\mu\nu}$, field equations and commutation relations:

$$t_{\mu\nu} = \frac{1}{4}(\partial_{(\mu} \gamma_{\lambda\rho} \cdot \partial_{\nu)} \gamma_{\lambda\rho} - \frac{1}{2}\partial_{(\nu} \gamma \cdot \partial_{\mu)} \gamma) - \frac{1}{4}\partial_{\mu\nu} (\partial_\sigma \gamma_{\lambda\rho} \cdot \partial_\sigma \gamma_{\lambda\rho} - \frac{1}{2}\partial_\lambda \gamma \cdot \partial_\lambda \gamma), \quad (2.2)^{**}$$

$$\square \gamma_{\mu\nu} = 0, \quad \square \gamma = 0, \quad (2.3)$$

$$[\gamma_{\mu\nu}(x), \gamma_{\lambda\rho}(x')] = i\delta_{(\mu\lambda} \delta_{\nu)\rho} D(x-x'),$$

$$[\gamma(x), \gamma(x')] = -4iD(x-x'). \quad (2.4)$$

Since the conditions $\partial_\nu \gamma_{\mu\nu} = 0$ and $\gamma_{\mu\mu} - \gamma = 0$ are not consistent with (2.4), they should be replaced by

$$\partial_\nu \gamma_{\mu\nu} \phi = 0, \quad (2.5)$$

$$(\gamma_{\mu\mu} - \gamma) \phi = 0 \quad (2.6)$$

respectively.*** We shall assume that the above conditions also hold when there are interactions. These supplementary conditions are used to exclude the unphysical parts of $\gamma_{\mu\nu}$ which gives rise to the negative energy terms. In other words, by these condition

* We take the space-time coordinates as x_1, x_2, x_3 and $x_4 = it$.

† In the linear approximation, the following relation between $\gamma_{\mu\nu}$ and metric tensor $g_{\mu\nu}$ follows from (1.1):

$$g_{\mu\nu} = \delta_{\mu\nu} + \kappa h_{\mu\nu}, \quad \gamma_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu} \cdot h, \quad h = h_{\mu\mu}.$$

†† In the following, we denote $A_{(\mu,\nu)} = A_{\mu,\nu} + A_{\nu,\mu}$, $A_{[\mu,\nu]} = A_{\mu,\nu} - A_{\nu,\mu}$.

*** Strictly, (2.5) is replaced by the following condition, as Gupta did,

$$[\partial_\nu \gamma_{\mu\nu}]^{(+)} \phi = 0$$

where $[\partial_\nu \gamma_{\mu\nu}]^{(+)}$ is the positive frequency part of $\partial_\nu \gamma_{\mu\nu}$. But for such a problem as to obtain the gravitational potential by eliminating the supplementary condition, no essential difference appears. For instance in the case of quantum electrodynamics, see, K. Bleuer, Helv. Phys. Acta 23 (1950), 5

the canonical energy-momentum tensor $t_{\mu\nu}$ may be considered as that of particle of spin 2, containing no particles of spin 0 and 1. As shown in § 4, only the transverse part of gravitational field $\bar{\gamma}_{\mu\nu}$ satisfies $\partial_\nu \bar{\gamma}_{\mu\nu} = 0$ rigorously, without use of operator equation such as (2.5), and so has the properties of particle in the strict sense. From the relation, $[\gamma_{\mu\mu}(x), \gamma_{\nu\nu}(x')] = 8i D(x-x')$, it appears that the supplementary condition (2.6) is not compatible with (2.4). But, this inconsistency is attributed to the fact that the first supplementary condition (2.5) is not taken into account. In fact, if we eliminate (2.5) completely, the compatibility of (2.6) and (2.4) is guaranteed, as shown in § 4.

As the first approximation, a pure gravitational field of spin 2 will be described by the field equation

$$\square \gamma_{\mu\nu} = \kappa t_{\mu\nu}. \quad (2.3')$$

According to the present ideas of the field theory, the field equation (2.3') must be derived from a Lagrangian density. We must, therefore, add the Lagrangian density L_G' to zeroth order Lagrangian density L_0 in such a way that in the resulting field equation the quantity (2.2) appears in the right-hand side of (2.3). In this way Gupta⁽⁴⁾ introduced an infinite number of terms in the Lagrangian density by successive approximations.

When there is a matter field, the problem becomes very complicated and further it seems that there exists no definite rule to obtain the interaction Lagrangian density. Therefore, we shall here introduce the interaction Lagrangian density in such a way as to be compatible with the supplementary condition (2.5), as follows:

We take as the starting Lagrangian density

$$L = L_G + L_G' + \dots + L_m - (\kappa/2) h_{\mu\nu} \cdot T_{\mu\nu} \quad (2.7)$$

where $h_{\mu\nu} = \gamma_{\mu\nu} - \frac{1}{2} \partial_{\mu\nu} \gamma$ and L_m means free Lagrangian density for the matter field, $T_{\mu\nu}$ energy-momentum tensor for the matter field derived from L_m .

As a simple case, we shall consider the scalar field $V(x)$ with spin 0 and mass m as a matter field, then,

$$L_m = -\frac{1}{2} [(\partial_\lambda V)^2 + m^2 V^2], \quad (2.8)$$

$$T_{\mu\nu} = \frac{1}{2} [\partial_{(\mu} V \cdot \partial_{\nu)} V - \partial_{\mu\nu} \{(\partial_\lambda V)^2 + m^2 V^2\}]. \quad (2.9)$$

The field equations are

$$\begin{aligned} \square \gamma_{\mu\nu} &= \kappa (T_{\mu\nu} + t_{\mu\nu}), \quad \square \gamma = \kappa (T + t), \\ (\square - m^2) V &= -\kappa [\partial_\nu (\partial_\mu V \cdot h_{\mu\nu}) - \partial_\mu (\partial_\nu V \cdot h/2) + m^2 V \cdot h/2]. \end{aligned} \quad (2.10)$$

Since $\partial_\nu (T_{\mu\nu} + t_{\mu\nu}) \neq 0$, the first equation of (2.10) is not compatible with the supplementary condition (2.5). At first sight, it seems that the compatibility will be satisfied by introducing an additive Lagrangian density in such a way that the following total canonical energy-momentum tensor $T'_{\mu\nu}$ appears in the right-hand side of the first equation of (2.10):

$$T'_{\mu\nu} = t_{\mu\nu} + T_{\mu\nu} + \kappa h_{\nu\lambda} \partial_\lambda V \cdot \partial_\mu V - (\kappa/2) h \cdot \partial_\mu V \cdot \partial_\nu V \\ - (\kappa/2) \partial_{\mu\nu} h_{\lambda\rho} [\partial_\lambda V \cdot \partial_\rho V - \frac{1}{2} \partial_{\lambda\rho} \{ (\partial_\sigma V)^2 + m^2 V^2 \}]. \quad (2 \cdot 11)$$

But, unfortunately, as $T'_{\mu\nu}$ is not symmetrical with μ and ν and $\partial_\mu T'_{\mu\nu} \neq 0$ ($\partial_\nu T'_{\mu\nu} = 0$), the field equation contradicts with (2.5). Therefore, in order to symmetrize $T'_{\mu\nu}$ in such a way that the modified energy-momentum tensor is divergenceless, we must use, from the beginning, the *real* energy-momentum tensor $t_{\mu\nu}$ for the gravitational field, instead of the canonical energy-momentum tensor $t_{\mu\nu}$. $t_{\mu\nu}$ is obtained from $t_{\mu\nu}$ by adding a divergenceless tensor ($\partial_\nu (t_{\mu\nu} - t_{\mu\nu}) = 0$, but $\partial_\nu (t_{\mu\nu} - t_{\mu\nu}) \neq 0$) which may be interpreted as the spin-energy-momentum tensor of graviton:

$$\bar{t}_{\mu\nu} = t_{\mu\nu} - \frac{1}{2} [\square \gamma_{[\nu\lambda} \cdot \gamma_{\lambda\mu]} + \partial_{(\mu} \gamma_{\rho\lambda} \cdot \partial_\rho \gamma_{\lambda\nu)} - \partial_\rho \partial_{(\mu} \gamma_{\nu)\lambda} \cdot \gamma_{\lambda\rho}]. \quad (2 \cdot 12)$$

When the matter field exists, the second term plays the role of symmetrizing the interaction term $\kappa h_{\nu\lambda} \cdot \partial_\lambda V \cdot \partial_\mu V$ in the form $(\kappa/2) h_{(\nu\lambda} \partial_\lambda V \cdot \partial_\mu V$. Therefore, when there are interactions including the non-linear part of the gravitational field, we have to introduce L'_i, L''_i, \dots so that the real energy-momentum tensor $t_{\mu\nu}$ given by (2.12) appears in the right-hand side of (2.3'), though Gupta⁴⁾ used $t_{\mu\nu}$ instead of $t_{\mu\nu}$.

If the above mentioned procedure is taken, we can obtain as the additive interaction Lagrangian density

$$L'_i = -(\kappa^2/4) [\{ (h_{\mu\lambda} - \frac{1}{2} \partial_{\mu\lambda} h) \partial_\lambda V \}^2 + (m^2/4) h^2 V^2]. \quad (2 \cdot 13)$$

In this way, additive Lagrangian densities L'_i, L''_i, \dots are introduced by successive approximations. However, it will be a mathematically complicated problem to obtain the higher order interaction terms than κ^3 .

§ 3. Interaction representation

For practical purposes it is most convenient to use the invariant perturbation theory based on the interaction representation. For interacting systems of the graviton and matter fields, there is, in general, no unitary transformation which connects the interaction representation and the Heisenberg representation. We shall, therefore, construct the interaction representation by the method which is the extension of the Yang-Feldman's⁵⁾ discussed by Takahashi, Umezawa and Katayama.⁶⁾

In the following, we shall take as the matter fields a Fermion of spin 1/2 and a Boson of spin 0 or 1, and we shall obtain the interaction Hamiltonian for each case. As we are concerned with the discussion of the interaction between the gravitational field and matter fields, we shall take the Lagrangian density (2.1') as that of a pure gravitational field. (Owing to this, there must appear some unsatisfactory features. They can be removed, however, by considering the non-linear part.)

i) Graviton-Boson (by wave aspect) system

Starting with the Lagrangian density (2.13), we replace $\gamma_{\mu\nu}(x, \sigma)$, $V(x, \sigma)$ etc., which are solutions of free equations and are connected with the freely oscillating incoming

fields $\gamma_{\mu\nu}^{\prime\prime\prime}$, $V^{\prime\prime\prime}$ by a unitary transformation function $U(\sigma, \sigma' = -\sigma\sigma')$, with the Heisenberg operators. The unitary transformation function can be represented by a Hermitian operator $H(x/\sigma)$,* which is nothing but the interaction Hamiltonian in the interaction representation, in the following form

$$U(\sigma, \sigma') = 1 - i \int_{\sigma'}^{\sigma} H(x'/\sigma) (dx') + \dots$$

in which $H(x'/\sigma)$ satisfies the following equations

$$\begin{aligned} -i[\gamma_{\mu\nu}(x, \sigma), H(x'/\sigma)] &= \kappa D(x-x') \mathbf{T}_{\mu\nu}(x'), \\ -i[\gamma(x, \sigma), H(x'/\sigma)] &= \kappa D(x-x') \mathbf{T}(x'), \\ i[V(x, \sigma), H(x'/\sigma)] &= \kappa [\partial_\nu D(x-x') \partial_\mu V(x') (\mathbf{h}_{\mu\nu}(x') - (\partial_{\mu\nu}/2) \mathbf{h}(x')) \\ &\quad + \frac{1}{2} m^2 D(x-x') V(x') \mathbf{h}(x')]. \end{aligned} \quad (3.1)$$

Expressing the Heisenberg operators** $\gamma_{\mu\nu}(x)$, $\gamma(x)$, $V(x)$ in the right-hand side of (3.1) by the quantities of x/σ , we can obtain the interaction Hamiltonian

$$\begin{aligned} H &= (\kappa/2) h_{\mu\nu} \partial_\mu V \cdot \partial_\nu V - (\kappa/4) h [(\partial_\lambda V)^2 + m^2 V^2] \\ &\quad + \frac{\kappa^2}{2} \frac{1}{1 - \kappa(\bar{h} + h/2)} [n_\lambda (\mathbf{h}_{\lambda\rho} - \frac{1}{2} \partial_{\lambda\rho} h) \partial_\rho V]^2, \end{aligned} \quad (3.2)$$

where we have simply written $V(x/\sigma)$, $\gamma_{\mu\nu}(x/\sigma)$ as V , $\gamma_{\mu\nu}$ respectively, and $\bar{h} = n_\lambda n_\rho h_{\lambda\rho}$. Though the interaction Hamiltonian does not close and has infinite terms of the coupling constant, we are able to express it in a unified form. It should be noted that the infinite series of normal dependent terms is caused by the interaction terms which contain γ and $\gamma_{\mu\nu}$ for which $n_\nu \gamma_{\mu\nu} \neq 0$ (say, the gravitational fields other than the transverse part, § 4).

ii) Graviton-Boson (by particle aspect) system

In order to discuss both a spin 0 particle and a spin 1 particle including the photon by the same type wave equation, it is convenient to describe the boson by Kemmer's β -formalism. The Lagrangian density for this system is given by

$$\begin{aligned} L &= -\frac{1}{4} (\partial_\lambda \gamma_{\mu\nu} \cdot \partial_\lambda \gamma_{\mu\nu} - \frac{1}{2} \partial_\lambda \gamma \cdot \partial_\lambda \gamma) - \frac{1}{2} (\bar{\phi} \beta_\mu \partial_\mu \phi - \partial_\mu \bar{\phi} \beta_\mu \phi + 2m \phi \bar{\phi}) \\ &\quad - (\kappa/2) \mathbf{h}_{\mu\nu} \Theta_{\mu\nu}, \end{aligned} \quad (3.3)$$

where

$$\Theta_{\mu\nu} = -m \bar{\phi} [\beta_{(\mu} \beta_{\nu)} - \delta_{\mu\nu}] \phi \quad (3.4)$$

and ϕ satisfies the commutation relation***

* x/σ means x laid on σ .

** In this section, the bold-faced type such as \mathbf{V} means the operator in the Heisenberg representation

*** When we take photon as the matter field, m must be replaced by a singular idempotent operator ω introduced by Harish-Chandra⁽¹⁾

$$\omega^2 = \omega, \quad \omega \beta_\mu + \beta_\mu \omega = \beta_\mu$$

and $S(x-x')$ by $S(x-x') = [\square + (\beta\partial) - (\beta\partial)^2] \omega \delta(x-x')$.

$$[\phi(x), \bar{\phi}(x')] = -iS(x-x') = -i\{(1/m)(\square - m^2) + (\beta\partial) - (1/m)(\beta\partial)^2\}A(x-x') \quad (3.5)$$

when the interaction is absent.

The interaction Hamiltonian becomes

$$\begin{aligned} H &= -\kappa m \bar{\phi}(\beta h)\phi - \kappa^2 \bar{\phi}(\beta h)[1 + (\beta n)^2](\beta h)\dot{\phi} + \dots \\ &= -\kappa m \bar{\phi}(\beta h) \frac{1}{1 - (\kappa/m)[1 + (\beta n)^2](\beta h)} \phi \end{aligned} \quad (3.6)$$

where

$$(\beta h) = (\beta_\mu \beta_\nu - \frac{1}{2} \partial_{\mu\nu}) h_{\mu\nu}. \quad (3.7)$$

In the photon-meson system where the interaction term has the form $(\beta A) = \beta_\mu A_\mu$ (A_μ : electromagnetic potential), only the first two terms of the series (3.6) survive in virtue of the identity

$$[1 + (\beta n)^2] \beta_\mu [1 + (\beta n)^2] = 0.$$

However, in our graviton-meson system the interaction Hamiltonian does not close owing to the relation: $[1 + (\beta n)^2](\beta_\mu \beta_\nu - \frac{1}{2} \partial_{\mu\nu})[1 + (\beta n)^2] \neq 0$.

Finally, we shall compare the interaction Hamiltonian (3.2) obtained by wave formalism with (3.6) obtained by β -formalism. As shown in Appendix, the first term of (3.2) agrees with that of (3.6). The second term of (3.6) becomes

$$H_{p2} = (\kappa^2/2) [\{ (h_{\lambda\rho} - \frac{1}{2} \partial_{\lambda\rho} \cdot h) \partial_\rho V \}^2 + \{ n_\lambda (h_{\lambda\rho} - \frac{1}{2} \partial_{\lambda\rho} \cdot h) \partial_\rho V \}^2] \quad (3.8)$$

which does not agree with that of (3.2). At first sight it seems strange that we obtained two different Hamiltonians according to different descriptions, despite we started with the same Lagrangian density. We shall now show that this discrepancy is removed by taking into account the non-linear part of gravitational field.

In § 2, we have shown that we must add to the additive interaction Lagrangian density to preserve the compatibility of the commutation relation and the supplementary condition. If we take account of this fact, κ^2 order term of interaction Hamiltonian becomes in the wave formalism

$$\begin{aligned} H'_{w2} &= H_{w2} - L'_{w1} \\ &= (\kappa^2/4) [\{ (h_{\lambda\rho} - \frac{1}{2} \partial_{\lambda\rho} \cdot h) \partial_\rho V \}^2 + (m^2/4) h^2 V^2] + (\kappa^2/2) [n_\lambda (h_{\lambda\rho} - \frac{1}{2} \partial_{\lambda\rho} \cdot h) \partial_\rho V]^2 \end{aligned} \quad (3.9)$$

where H_{w2} is the second term of (3.2) and L'_{w1} means (2.13).

Also, in β -formalism, for the same reason we must add a term L'_{p1} to the original interaction Lagrangian density $\kappa m \bar{\phi}(\beta_\mu \beta_\nu - \frac{1}{2} \partial_{\mu\nu}) \phi \cdot h_{\mu\nu}$. It can be shown that the canonical energy-momentum tensor $T_{\mu\nu}$:

$$T_{\mu\nu} = \frac{1}{2} [\bar{\phi} \beta_\nu \partial_\mu \phi - \partial_\mu \bar{\phi} \beta_\nu \phi] \quad (3.10)$$

is related to the symmetrical energy-momentum tensor $\theta_{\mu\nu}$ by the relation:

$$\begin{aligned} T_{\mu\nu} = & \Theta_{\mu\nu} - \kappa \bar{\phi} [(\beta_\nu \beta_\mu - \frac{1}{2} \delta_{\mu\nu}) (\beta \cdot \mathbf{h}) + (\beta \mathbf{h}) (\beta_\mu \beta_\nu - \frac{1}{2} \delta_{\mu\nu})] \phi \\ & - \frac{1}{2} \partial_\rho [\phi (\beta_\rho \beta_\mu \beta_\nu - \beta_\nu \beta_\mu \beta_\rho) \phi]. \end{aligned} \quad (3 \cdot 10')$$

With the aid of this relation, the additive interaction term L'_{pI} can be obtained:

$$L'_{pI} = -(\kappa^2/2) \phi (\beta h)^2 \phi. \quad (3 \cdot 11)$$

Taking into account this term, the κ^2 order term of interaction Hamiltonian in the β -formalism becomes

$$H'_{pI} = H_{p2} - L'_{pI} = -\kappa^2 \bar{\phi} (\beta h) (\beta n)^2 (\beta h) \phi - (\kappa^2/2) \bar{\phi} (\beta h)^2 \phi \quad (3 \cdot 12)$$

where H_{p2} means the second term of (3.6). As shown in Appendix, (3.12) agrees with (3.9). Thus, the discrepancy of two Hamiltonians obtained in each formalism can be removed by considering the non-linear character of the gravitational field.

iii) Graviton-Fermion system

Lagrangian density is

$$\begin{aligned} L = & -\frac{1}{4} [\partial_\lambda \bar{\gamma}_{\mu\nu} \cdot \partial_\lambda \gamma_{\mu\nu} - \frac{1}{2} \partial_\lambda \bar{\gamma} \cdot \partial_\lambda \gamma] - \frac{1}{2} [\bar{\phi} \alpha_\mu \partial_\mu \phi - \partial_\mu \bar{\phi} \alpha_\mu \phi + 2m \bar{\phi} \phi] \\ & - (\kappa/4) \mathbf{h}_{\mu\nu} (\bar{\phi} \alpha_\mu \partial_\nu \phi - \partial_\mu \bar{\phi} \alpha_\mu \phi) \end{aligned} \quad (3 \cdot 13)$$

where α_μ is the Dirac matrix. The interaction term resembles the Konopinski-Uhlenbeck type one. Hamiltonian is determined by the following equations:

$$\begin{aligned} -i[\gamma_{\mu\nu}(x, \sigma), H(x'/\sigma)] &= (\kappa/2) D(x-x') [\bar{\phi}(x') \alpha_\mu \partial_\nu \phi(x') - \partial_\nu \bar{\phi}(x') \alpha_\mu \phi(x')], \\ i[\psi(x, \sigma), H(x'/\sigma)] &= (\kappa/4) [S(x-x') \mathbf{h}_{\mu\nu}(x') \alpha_\mu \partial_\nu' \psi(x') \\ &+ \partial_\nu S(x-x') \mathbf{h}_{\mu\nu}(x') \alpha_\mu \psi(x')] + F(x, x') \end{aligned} \quad (3 \cdot 14)$$

where $F(x, x')$ is determined so as to be able to have a solution. If we choose $F(x, x')$ as follows:

$$\begin{aligned} F(x, x') = & -(\kappa^2/32) \{S(x-x') (nA) (\alpha\partial) (nA) \cdot \psi + \partial_\lambda [S(x-x') (nA) \psi] \alpha_\lambda (nA) \psi \\ & + S(x-x') (nA) (n\alpha) \partial_\lambda (A_\lambda \psi) + \partial_\lambda [S(x-x') (nA)] (n\alpha) A_\lambda \psi \\ & + S(x-x') A_\lambda (n\alpha) \partial_\lambda [(nA) \psi] + \partial_\lambda [S(x-x') A_\lambda] (n\alpha) (nA) \psi \\ & + 2S(x-x') (nA) (n\alpha) (n\partial) [(nA) \psi] \\ & - 2(n\partial) [S(x-x') (nA)] (n\alpha) (nA) \psi\} + O(\kappa^3) \end{aligned} \quad (3 \cdot 15)$$

where $A_\mu = \mathbf{h}_{\mu\nu} \alpha_\nu$, we can obtain the interaction hamiltonian:

$$\begin{aligned} H = & (\kappa/4) \bar{\phi} [1 - (\kappa/4) (nA) (n\alpha)] [A_\nu + (nA) (n\alpha) \alpha_\nu] \partial_\nu^s \phi - \text{comp. conj.} \\ & + (\kappa^2/32) [\bar{\phi} (nA) (\alpha\partial^s) \{ (nA) \psi\} + \bar{\phi} (nA) (n\alpha) \partial_\lambda^s (A_\lambda \psi) + \bar{\phi} A_\lambda (n\alpha) \partial_\lambda^s \{ (nA) \psi\} \\ & - \bar{\phi} (nA) (n\alpha) (n\partial) (nA) \cdot \psi + \bar{\phi} (nA) (n\alpha) (nA) (n\alpha) (\alpha\partial^s) \psi] - \text{comp. conj.} \\ & + (\kappa/2) m \bar{\phi} [\bar{h} - (\kappa/4) \{\bar{h}^2 + (n_\lambda h_{\lambda\rho})^2\} \psi + O(\kappa^3) (\partial_\mu^s = \partial_\mu + n_\mu (n\partial))]. \end{aligned} \quad (3 \cdot 16)$$

Different from the cases i) and ii), Hamiltonian cannot be written in a simple form.

iv) *Supplementary condition*

In § 2, we took the supplementary conditions (2.5) and (2.6) in the Heisenberg representation. Now, we shall investigate in what forms these conditions can be written in the interaction representation. (2.6) has the same form as in the Heisenberg representation, with a proviso that ϕ is replaced by a state vector $\Psi(\sigma)$ and x_μ is not necessarily restricted on the surface σ . As an example, we consider the system of graviton and meson described by β -formalism.

The following equations being taken into account

$$-i[\gamma_{\mu\nu}(x), H(x')] = \kappa m D(x-x') \phi(x') \frac{1}{1 - (\kappa/m)(\beta h') [1 + (\beta n)^2]} [\beta_{(\mu} \beta_{\nu)} - \delta_{\mu\nu}] \\ \times \frac{1}{1 - (\kappa/m) [1 + (\beta n)^2] (\beta h')} \phi(x')$$

and $(\beta n)[1 + (\beta n)^2] = 0$, (2.5) becomes

$$[\partial_\nu \gamma_{\mu\nu} + \kappa \int D(x-x') \{ \theta_{\mu\nu}(x') - \kappa \phi(x') [(\beta_\nu \beta_\mu - \delta_{\mu\nu})(1 + (\beta n)^2)(\beta h') \\ + \dots] \phi(x') \} d\sigma_\nu'] \Psi(\sigma) = 0 \quad (3.17)$$

where $\theta_{\mu\nu}$ is the symmetrical energy-momentum tensor of meson in the interaction representation.

On the other hand, in order to satisfy the compatibility condition of the commutation relation and the supplementary condition, the second term of (3.10') is necessary in addition to $\theta_{\mu\nu}$. Considering this situation and the identity $[(\beta n)_i \beta_\mu - n_\mu](\beta n)^2 = 0$, (3.17) reduces, up to κ^2 order term, to

$$[\partial_\nu \gamma_{\mu\nu} + \kappa \int D(x-x') \{ \theta'_{\mu\nu} + \kappa \partial_{\mu\nu}(\beta h') \} d\sigma_\nu] \Psi(\sigma) = 0 \quad (3.18)$$

where x_μ is not necessarily restricted on the surface σ . As (3.17) shows, the normal dependent terms do not drop in the supplementary condition contrary to the case of the usual method. However, by considering the non-linear character of gravitational field, we can show that the normal dependent terms drop and $-\partial_{\mu\nu} L_I$ appears in place of the third term of (3.17).

Finally, in connection with (3.18), we shall recall the following relation between the energy-momentum tensor $\theta_{\mu\nu}$ in the Heisenberg representation and the free field one, $\theta_{\mu\nu}$, in the interaction representation:⁹⁾

$$\theta_{\mu\nu} - \partial_{\mu\nu} H(x, n) = U(\sigma) \theta_{\mu\nu} U^{-1}(\sigma). \quad (3.19)$$

In our subsidiary condition, only the normal independent terms in the left-hand side of (3.19) appear by considering the additive interaction Lagrangian densities L_I, L_I'', \dots .

§ 4. Elimination of supplementary condition and gravitational potential

As shown in § 3 iv), the supplementary condition in the interaction representation is

$$[\partial_\nu \gamma_{\mu\nu}(x) + \kappa \int D(x-x') \underline{T}_{\mu\rho}(x') d\sigma_\rho] \Psi(\sigma) = 0 \quad (4.1)$$

where $\underline{T}_{\mu\nu}$ is related to the energy-momentum tensor for the free field $T_{\mu\nu}$ and the interaction Lagrangian L_I by the relation

$$\underline{T}_{\mu\nu} = T_{\mu\nu} + \delta_{\mu\nu} L_I. \quad (4.2)$$

Here, for brevity we do not take into account the second term of (4.2) whose appearance is due to the non-linear character of the gravitational field as mentioned in § 3 iv). Then

$$[\partial_\nu \gamma_{\mu\nu} + \kappa \int D(x-x') T_{\mu\rho}(x') d\sigma_\rho] \Psi = 0. \quad (4.1')$$

We shall now eliminate the gravitons other than the transverse one from the equation of motion by making use of the supplementary condition (4.1'), and shall obtain a gravitational potential between the matters.

For this purpose, we decompose $\gamma_{\mu\nu}(x)$ into the scalar fields $A^{(\alpha\beta)}(x)$ ($\alpha, \beta = 1, 2, 3$), as

$$\gamma_{\mu\nu}(x) = N_\mu^{(\alpha)} N_\nu^{(\beta)} A^{(\alpha\beta)}(x) \quad (4.3)$$

where $N_\mu^{(\alpha)}$ are defined by

$$\begin{aligned} N_\mu^{(4)} &= n_\mu, \quad N_\mu^{(3)} = [\partial_\mu + n_\mu (h\partial)] / (n\partial), \\ N_\mu^{(\alpha)} N_\mu^{(\beta)} &= -\varepsilon_{\alpha\beta}, \quad N_\mu^{(\alpha)} N_\nu^{(\beta)} \varepsilon_{\alpha\beta} = -\partial_{\mu\nu}, \quad N_\nu^{(1)} \partial_\nu = N_\nu^{(2)} \partial_\nu = 0, \\ (\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{33} = -\varepsilon_{44} = -1; \quad \varepsilon_{\mu\nu} &= 0 \quad (\mu \neq \nu)) \end{aligned} \quad (4.4)$$

We define the pure transverse part of field quantities by

$$\bar{\gamma}_{\mu\nu}(x) = \sum_{\alpha, \beta=1}^2 N_\mu^{(\alpha)} N_\nu^{(\beta)} A^{(\alpha\beta)}(x) \quad (4.5)$$

which satisfies

$$n_\nu \bar{\gamma}_{\mu\nu} = \partial_\nu \bar{\gamma}_{\mu\nu} = 0. \quad (4.6)$$

Above equations assure that the transverse part has particle characters. From the commutation relation (2.4) and equations (4.3) and (4.4), we obtain the commutation relation for $A^{(\alpha\beta)}(x)$:

$$[A^{(\alpha\beta)}(x), A^{(\gamma\delta)}(x')] = i \varepsilon_{(\alpha\gamma} \varepsilon_{\beta\delta)} D(x-x'). \quad (4.7)$$

Further, for the convenience of later treatments we introduce $A_\mu(x)$ and $\bar{A}_\mu(x)$ defined by

$$A_\mu(x) = N_\mu^{(\alpha)} A^{(\alpha 4)}(x), \quad (4.8)$$

$$\bar{A}_\mu(x) = N_\mu^{(\alpha)} A^{(\alpha 3)}(x). \quad (4.9)$$

Then we can show that $A_\mu(x)$, $\bar{A}_\nu(x)$ and $\bar{\gamma}_{\mu\nu}(x)$ satisfy the following commutation relations:

$$[\bar{\gamma}_{\mu\nu}(x), A_\sigma(x')] = [\bar{\gamma}_{\mu\nu}(x), \bar{A}_\sigma(x')] = 0 \quad (4.10)$$

and

$$[\tilde{\gamma}_{\mu\nu}(x), \tilde{\gamma}_{\lambda\rho}(x')] = i[\partial_{(\mu}^s \partial_{\nu)}^s - \partial_{(\mu}^s \partial_{\nu)}^s \partial_{\rho}^s (n\partial)^{-2} - \partial_{(\mu}^s \partial_{\nu)}^s \partial_{\lambda}^s (n\partial)^{-2} + 2\partial_{\mu}^s \partial_{\nu}^s \partial_{\lambda}^s \partial_{\rho}^s (n\partial)^{-4}] D(x-x'). \quad (4.11)$$

where

$$\partial_{\mu\nu}^s = \partial_{\mu\nu} + n_{\mu} n_{\nu}, \quad \partial_{\mu}^s = \partial_{\mu} + n_{\mu} (n\partial). \quad (4.12)$$

(4.11) is, of course, compatible with the restriction (4.6).

The interaction Hamiltonian is expressed by $\tilde{\gamma}_{\mu\nu}(x)$, $A_{\mu}(x)$ and $\bar{A}_{\nu}(x)$ in the following form

$$\begin{aligned} H = & (\kappa/2) \tilde{\gamma}_{\mu\nu} T_{\mu\nu} - (\kappa/4) \gamma \cdot T + (\kappa/2) (N_{\mu}^{(3)} N_{\nu}^{(1)} A^{(31)} + N_{\mu}^{(1)} N_{\nu}^{(2)} A^{(32)} \\ & + N_{\mu}^{(4)} N_{\nu}^{(1)} A^{(41)} + N_{\mu}^{(1)} N_{\nu}^{(2)} A^{(42)}) T_{\mu\nu} \\ & + (\kappa/2) N_{\nu}^{(3)} (A_{\mu} + \bar{A}_{\mu}) T_{\mu\nu} - (\kappa/2) \partial_{\nu} [(n\partial)^{-1} A_{\mu}] \cdot T_{\mu\nu}. \end{aligned} \quad (4.13)$$

The supplementary condition (4.1') can be written as

$$[(n\partial) (A_{\mu}(x) + \bar{A}_{\mu}(x)) + \kappa \int D(x-x') T_{\mu\rho}(x') d\sigma_{\rho}'] \Psi = 0. \quad (4.14)$$

We introduce the canonical transformation:

$$\Psi \rightarrow e^{-iG} \Psi \quad (4.15)$$

where

$$\begin{aligned} G = & (\kappa/2) \int (n\partial')^{-1} [2 (N_{\lambda}^{(1)} A^{(31)}(x') + N_{\lambda}^{(2)} A^{(32)}(x') + N_{\lambda}^{(3)} A^{(33)}(x') \\ & - N_{\lambda}^{(4)} A^{(41)}(x')) T_{\lambda\rho}(x') d\sigma_{\rho}']. \end{aligned} \quad (4.16)$$

in which $A^{(31)}$, $A^{(32)}$, $A^{(33)}$ and $A^{(41)}$ are all Hermitian operators. The supplementary condition for the new state vector becomes

$$[A_{\mu}(x) + \bar{A}_{\mu}(x)] \Psi(\sigma) = 0. \quad (4.17)$$

Further, multiplying $N_{\nu}^{(1)}$, $N_{\nu}^{(2)}$, $N_{\nu}^{(3)}$ and $N_{\nu}^{(4)}$, and carrying out the summation over μ , (4.17) becomes

$$[A^{(13)} + A^{(14)}] \Psi = [A^{(23)} + A^{(24)}] \Psi = [A^{(33)} + A^{(34)}] \Psi = [A^{(43)} + A^{(44)}] \Psi = 0. \quad (4.18)$$

Under the transformation (4.15), the interaction Hamiltonian (4.13), in view of the supplementary condition, is transformed into

$$\begin{aligned} H = & (\kappa/2) \tilde{\gamma}_{\mu\nu} T_{\mu\nu} - (\kappa/4) \gamma \cdot T + (\kappa^2/4) \int (n\partial)^{-1} D(x-x') n_{\mu} n_{\nu} n_{\lambda} T_{\mu\nu}(x) T_{\lambda\rho}(x') d\sigma_{\rho}' \\ & - (\kappa^2/2) \int (n\partial)^{-1} D(x-x') n_{\mu} \partial_{\nu\lambda}^s T_{\mu\nu}(x) T_{\lambda\rho}(x') d\sigma_{\rho}' \\ & + (\kappa^2/4) \int \partial_{\nu}^s \partial_{\lambda}^s (n\partial)^{-3} D(x-x') n_{\mu} T_{\mu\nu}(x) T_{\lambda\rho}(x') d\sigma_{\rho}' \end{aligned} \quad (4.19)$$

after some calculation. We have thus succeeded in constructing an equation of motion for $\mathcal{F}(\sigma)$ which no longer contains the gravitational field variables involved in the supplementary condition (4.1').

We denote the spatial integral of the sum of the third, fourth and fifth terms in (4.19) by V' . In view of the identities

$$(n\partial)^{-1}D(x-x') = 1/4\pi r, \quad r = \sqrt{(x_\mu - x'_\mu)^2}, \quad n_\mu(x_\mu - x'_\mu) = 0 \quad (4.20)$$

and

$$(n\partial)^{-3}D(x-x') = (1/8\pi)r, \quad n_\mu(x_\mu - x'_\mu) = 0 \quad (4.21)$$

which can be derived from the identity

$$\Delta(n\partial)^{-3}D(x-x') = (n\partial)^2(n\partial)^{-2}(n\partial)^{-1}D(x-x') = 1/4\pi r,$$

V' is expressed, in the special reference system $n_\mu = (0, 0, 0, i)$, as

$$\begin{aligned} V' = G \iint (d^3x)(d^3x') [-T_{44}(x)(1/r)T_{44}(x') - 2T_{4i}(x)(1/r)T_{4i}(x') \\ + \frac{1}{2}\partial^i\partial^j(r) \cdot T_{il}(x)T_{lj}(x')] \end{aligned} \quad (4.22)$$

where G is the Newton's gravitational constant:

$$G = (\kappa/2)^2/4\pi. \quad (4.23)$$

When one of the two interacting matters is photon (for which $T_{\mu\mu} = 0$) and the other matter is at rest, the gravitational potential between them is given by the first term of (4.22) which is two times larger than that of Newton's gravitational theory. This fact was useful to predict correctly the deflection of light rays which pass near the surface of sun.

In the following, for the sake of simplicity, we consider the case where two scalar particles with spin 0 are interacting through the gravitons. In the configuration space, $T_{\mu\nu}$ and T are, then,

$$\begin{aligned} T_{44} : \sqrt{m_1^2 + p_1^2} + \sqrt{m_2^2 + p_2^2}, \quad T_{4j} : -i(p_1^j + p_2^j) \\ T : m_1/\sqrt{m_1^2 + p_1^2} + m_2/\sqrt{m_2^2 + p_2^2}. \end{aligned} \quad (4.24)$$

Further, we shall now use the approximation method which consists in neglecting terms more than quadratic in the velocities of matters, as Einstein, Infeld and Hoffman did.

In addition to V' , there is the potential energy V^r due to the interaction $-(\kappa/4)\gamma \cdot T$ which is the first kind interaction (see, § 5):

$$V^r = \frac{G}{2} \iint (d^3x)(d^3x') T(x) \frac{1}{r} T(x') + V_{\text{ret}}^r. \quad (4.25)$$

The second term, say V_{ret}^r , means retarded effect and is expressed as¹¹⁾

$$V_{\text{ret}}^r = -\frac{G}{4} \iint (d^3x)(d^3x') \frac{d}{dt} \frac{d}{dt'} r \cdot T(x) \cdot T(x'). \quad (4.26)$$

With the aid of the relation

$$d/dt_j = i[H_j^0, \quad]$$

where H_j^0 is the free Hamiltonian for j -th particle, (4.26) can be written, in the configuration space, as

$$V_{\text{ret}}^{\text{ret}} = \frac{G}{2} m_1 m_2 [H_1^0, [H_2^0, r]] = -\frac{G}{8} [p_1^2, [p_2^2, r]], \quad r = \sqrt{(x_1 - x_2)^2}. \quad (4.26')$$

In view of (4.24) and (4.26'), the last term of (4.22), say $V_{\text{ret}}' = (G/2) \iint (d^3x) \times (d^3x') \partial^i \partial^j (r) T_{4i}(x) T_{4j}(x')$, is equal to $-2V_{\text{ret}}^{\text{ret}}$. We transform total $V_{\text{ret}} = V_{\text{ret}}^{\text{ret}} + V_{\text{ret}}'$ $= -V_{\text{ret}}^{\text{ret}}$ so as to be interpreted as the potential term appeared in Dirac's many time formalism. By the transformation independent of past history:

$$\phi = \exp[(i/2) G m_1 m_2 (d/dt_1 + d/dt_2) r] \phi' \quad (4.27)$$

V_{ret} is modified as

$$\bar{V}_{\text{ret}} = \frac{G}{2} m_1 m_2 [p_2^2/2m_2, [p_2^2/2m_2, r]] + \frac{G^2}{2} (m_1 m_2)^2 [[p_1^2/2m_1 + p_2^2/2m_2, r], 1/r] \quad (4.28)$$

in which the last term has resulted from the correction for Newton potential by the transformation (4.27). (4.28) reduces to

$$\bar{V}_{\text{ret}} = -\frac{G}{2} \frac{m_2}{m_1} \left[p_2^2 \frac{1}{r} - p_2^2 - (p_2 r) - \frac{1}{r^3} (r p_2) \right] + \frac{G^2}{2} m_1 m_2 (m_1 + m_2) / r^2 \quad (4.28')$$

where we symmetrize the terms so as to preserve the Hermitean property manifestly. Thus, the effective potential is written, in the configuration space, as

$$V = -G m_2 \left[m_1 \frac{1}{r} + \frac{3}{2} \frac{1}{m_1} p_1^2 \frac{1}{r} - p_1^2 + 2 \frac{m_1}{m_2} p_2^2 \frac{1}{r} - p_2^2 - 2 \frac{1}{m_2} \left(p_1^2 \frac{1}{r} - p_2^2 + p_2^2 \frac{1}{r} - p_1^2 \right) \right. \\ \left. - \frac{1}{2} \frac{m_1}{m_2^2} (p_2 r) \frac{1}{r^3} (r p_2) - \frac{G}{2} m_1 \left(\frac{m_1}{r^2} + \frac{m_2}{r^2} \right) \right]. \quad (4.29)$$

There still remains the effect of the transverse part of graviton. As the interaction between the transverse graviton and scalar particle is of the second kind (§ 5) for which the validity of the perturbation expansion is doubtful, the potential obtained by the process of exchange of one graviton will have no definite meaning. Even if the perturbation expansion is admitted, the transverse graviton has not an effect within the scope of the approximation procedure of Einstein, Infeld and Hoffmann, as seen from the commutation relations for $\gamma_{\mu\nu}(x)$ given by (4.11).⁽¹²⁾ Thence, we should derive the potential without use of the transverse graviton. That is, for the two-body problem we should use the potential (4.29) derived by eliminating the supplementary condition. Using our potential defined by (4.29), we evaluate the second time derivative of the coordinate r_i^i by the formula

$$r_1^i = -[H, [H, r_1^i]] \quad (4.30)$$

where

$$H = p_1^2/2m_1 + p_2^2/2m_2 - (p_1^2)^2/8m_1^3 - (p_2^2)^2/8m_2^3 + \dots + V$$

(4.30) gives the equation of motion :

$$\begin{aligned} \ddot{r}_1^i = G \left\{ \varphi_{,i} + \frac{1}{m_1^2} p_1^s \varphi_{,s} p_1^i - 2 \frac{1}{m_1^2} (p_1^s \varphi_{,s} p_1^i + p_1^i \varphi_{,s} p_1^s) - 2 \frac{1}{m_1 m_2} (p_1^s \varphi_{,s} p_2^i + p_2^s \varphi_{,s} p_1^i) + \right. \\ \left. + 2 \frac{1}{m_1 m_2} (p_1^s \varphi_{,s} p_2^i + p_2^s \varphi_{,s} p_1^i) + \frac{3}{2} \frac{1}{m_1 m_2} (p_1^i \varphi_{,s} p_2^s + p_2^s \varphi_{,s} p_1^i) + 2 \frac{1}{m_2^2} p_2^s \varphi_{,s} p_2^i \right. \\ \left. - 2 \frac{1}{m_2^2} (p_2^i \varphi_{,s} p_2^s + p_2^s \varphi_{,s} p_2^i) - \frac{1}{2m_2} \partial_1^i \left[(p_2^s r) \frac{1}{r^3} - (r p_2^s) \right] \right\} - G^2 \varphi_{,i} \left(\frac{4m_2}{r} + \frac{5m_1}{r} \right) \end{aligned} \quad (4.31)$$

where $\varphi_{,i} = \partial_i(m_2/r)$. If we replace \mathbf{p} by $m\mathbf{v} = m\dot{\mathbf{x}}$, (4.31) is nothing but the equation of Einstein, Infeld and Hoffmann.¹⁾ Thus, we have shown that only the unquantizable part gives the effect on macroscopical phenomena.

Finally, we mention about the remark given in § 2. In view of (4.18), the supplementary condition (2.6) for the new state vector $\Psi(\sigma)$ is

$$[\bar{\gamma}_{\mu\mu} - \gamma - \kappa \int [\partial_\lambda + 2n_\lambda(n\partial)](n\partial)^{-2} D(x-x') \cdot T_{\lambda\rho}(x') d\sigma_\rho'] \Psi(\sigma) = 0. \quad (4.32)$$

From the commutation relation (4.11), we can easily show that $[\bar{\gamma}_{\mu\mu}(x), \bar{\gamma}_{\nu\nu}(x')] = 4iD(x-x')$ which no longer contradicts with the commutation relation for $\gamma(x)$.

§ 5. Divergence problems

As shown in § 3, the interaction Hamiltonian in the interaction representation does not close. This shows that some part of the interaction term between the gravitational field and matter field will have a strong field reaction and will give the large damping effect. Therefore we can expect that the non-singular theory of the interaction between elementary particles may be obtained if we appropriately take into account the large field reaction due to the interaction between the graviton and matter fields. In this section, we shall show that there is a possibility to obtain a convergent theory by using the propagation function of matter field modified by gravitational field.

In order to see whether the field reaction is so large as to suppress the high momenta states or not, we first investigate to what kind the interaction between the graviton and matter field belongs. For our purpose, it may be convenient to treat the problem in the Dyson's S -matrix formalism without separating the gravitational field into longitudinal, scalar and transverse parts. In the calculation of S -matrix, if we replace P -symbol by P^* -symbol, for instance, in the case of the scalar matter field

$$P^*(\partial_\mu V(x), \partial_\nu V(x')) = \frac{1}{2} \partial_\mu \partial_\nu \Delta_F(x-x'),$$

it is sufficient to deal only with the first term of the interaction Hamiltonian. Therefore,

for the calculation of S -matrix, we can consider that the graviton interacts with the matter fields via the following interaction terms

$$H = (\kappa/2) h_{\mu\nu} \partial_\mu V \partial_\nu V - (\kappa/4) h \{ (\partial_\lambda V)^2 + m^2 V^2 \} \quad \text{for matter field with spin 0,} \quad (5.1)$$

$$H = (\kappa/4) h_{\mu\nu} [\bar{\psi} \alpha_\mu \partial_\nu \psi - \partial_\nu \bar{\psi} \cdot \alpha_\mu \psi] \quad \text{for matter field with spin 1/2.} \quad (5.2)$$

Among the above terms, only the interaction terms of the scalar graviton γ and the matter fields :

$$H = (\kappa/4) m^2 V^2 \cdot \gamma \quad \text{for matter field with spin 0,} \quad (5.3)$$

$$H = (\kappa/4) m \bar{\psi} \psi \cdot \gamma \quad \text{for matter field with spin 1/2} \quad (5.4)$$

are the interactions of the first kind. The other interaction terms are of the second kind which cannot be renormalizable by the usual method based on the perturbation expansion. (But the interaction term which contains the longitudinal part of gravitational field can be eliminated with the aid of the supplementary condition, as shown in § 4.) In the previous section, we showed that the interaction of the first kind, for which the perturbation expansion is admitted, and the potential obtained by eliminating the longitudinal graviton have some effect upon the macroscopical phenomena. As the interaction term between the transverse graviton and matter field is of the second kind, it will give a large damping effect.

In the following, we use the method in which some parts of the large field reaction due to the interaction of the second kind are taken into account beforehand, and then consider the remaining parts. This method was originally proposed by N. Hu¹⁾ and generalized by Kamefuchi and Umezawa.¹¹⁾ They introduced the auxiliary C-meson field whose interaction with the electron is of the second kind. By using the internal electron line corrected by some part of the interaction of the second kind instead of the usual uncorrected one, they got the non-singular (convergent) quantum electrodynamics. But, in our case, instead of the artificial C meson field, we use the gravitational field which may be considered to exist in nature.

In order to obtain the propagation function of matter fields corrected by the gravitational field, we calculate the second order self-energy part $\Sigma(p)$ of matter field due to the graviton. In the first place, we shall consider the case where the matter field is the particle with spin 1/2. $\Sigma^F(p)$ is defined by the following part of S -matrix

$$S = -i \int_{-\infty}^{+\infty} (dx) \bar{\psi}(p) \Sigma^F(p) \psi(p) \exp(ipx). \quad (5.5)$$

By means of the usual Feynman-Dyson method, we find

$$\begin{aligned} \Sigma^F(p) = & (\kappa^2/128\pi^2) \left[\frac{3}{2} m \log(4/3) \cdot \xi^2 - 4m^2 \log(3\xi^2/8m^2) \right. \\ & + i \{ 9/2 \cdot \log(4/3) \cdot \xi^2 + 4m^2 \log(3\xi^2/8m^2) \} \\ & \times (\alpha p - im) + m \log(3\xi^2/8m^2) \cdot (\alpha p - im)^2 + i \log(3\xi^2/8m^2) \cdot (\alpha p - im)^3 \\ & \left. - ip^2 (\alpha p - im) \{ (1 - m^2/p^2) (1 - (m^2/p^2)^2) \log(|p^2 + m^2|/m^2) - \frac{5}{3} + 2m^2/p^2 - (m^2/p^2)^2 \} \right] \end{aligned}$$

$$-mp^2 \{ (3 + m^2/p^2) (1 - (m^2/p^2)^2) \log(|p^2 + m^2|/m^2) - 10/3 + 2m^2/p^2 + (m^2/p^2)^2 \}] \quad (5.6)$$

where the parameter ξ , which is of the dimension of mass, tends to infinity. Strictly speaking, (5.6) is correct only for $p^2 > -m^2$. For the case of $p^2 < -m^2$, logarithm appearing in the integral calculation by which (5.6) is obtained is not real. Accordingly, we must take the analytic continuation from large values of p to small values by writing $(p + i0)$ for p in the integral calculation, treating δ_0 as small and positive. Hence, for $p^2 < -m^2$, the following term is to be added to (5.6):

$$(\kappa^2/128\pi)\varepsilon(p_0)[p^2(\alpha p - im)(1 - m^2/p^2)(1 - (m^2/p^2)^2) - imp^2(3 + m^2/p^2)(1 - (m^2/p^2)^2)]. \quad (5.6')$$

The divergent term of $\Sigma^F(p)$ has the form

$$K_1 + K_2(\alpha p - im) + K_3(\alpha p - im)^2 + K_4(\alpha p - im)^3.$$

While the first and second terms are removed by renormalizing the mass and the coupling constant κ , third and fourth are not removed by the usual renormalization method.* Therefore, as a provisional method to remove the infinities, we add $\zeta^h[K_1 + K_2(\alpha p - im) + K_3(\alpha p - im)^2 + K_4(\alpha p - im)^3]\zeta^h$ to the Lagrangian density from the beginning. Then, we shall here leave out consideration for the divergent term appeared in $\Sigma^F(p)$.

We denote the finite part of $\Sigma^F(p)$ by $\Sigma_c^F(p)$. Taking into account all the iteration processes of $\Sigma_c^F(p)$, the propagation function of $S_F^F(p)$ becomes

$$S_F^{F'}(p) = S_F^F(p) / (1 - S_F^F(p) \Sigma_c^F(p)). \quad (5.7)$$

For $p^2 \gg m^2$, as seen from (5.6)

$$S_F^{F'}(p) \sim 1/p^3. \quad (5.8)$$

The corrected propagation function $S_F^{F'}(p)$ damps faster than the uncorrected one by $1/p^2$. Therefore, this causes sufficiently rapid decrease of the integrand at infinity.

As the gravitational field interacts with Fermion and Boson on an equal footing, we must also replace the free propagation function of Boson by the one modified by the field reaction due to the interaction between the graviton and Boson. In order to treat the scalar meson and vector meson simultaneously with the photon, we calculate by Kemmer's β -formalism. As the interaction term we take the first term of (3.6). By the same method as the case of the Fermion, the second order self-energy part $\Sigma^B(p)$ of Boson due to the gravitational field becomes, for $p^2 > -m^2$,

* Recently, Arnowitt & Deser and Cooper⁽⁵⁾ have shown that the divergence occurring in the interaction of the second kind has the form of essential singularities in contradistinction to the branch-point behavior of the usual "renormalizable" theory. Accordingly, the validity of the perturbation expansion is lost. By using the method which is not based on the perturbation expansion, they removed the divergent term occurring in the meson theories with derivative coupling. The infinities occurring in our system may be removed by the same method.

$$\begin{aligned}
\Sigma^B(p) = & (\kappa^2/16\pi^2) [-3(1/m) \log(4/3) \cdot \xi^2 B_3 + (1/m) (p^2/3 + m^2) B_3 \\
& + (1/m) (p^2 B_1 - (\beta p)^2 B_2) \{-\log(3\xi^2/8m^2) + \log(|p^2 + m^2|/m^2) \\
& - 11/6 + m^2/2p^2 - (m^2/p^2)^2 + (m^2/p^2)^3 \log(|p^2 + m^2|/m^2)\} \\
& + \{i(\beta p) B_4 - (1/m) (p^2 - m^2) B_5\} \{-\log(3\xi^2/8m^2) \\
& + \log(|p^2 + m^2|/m^2) - 3/2 + m^2/p^2 - (m^2/p^2)^2 \log(|p^2 + m^2|/m^2)\} \\
& + m B_5 \{-\log(3\xi^2/8m^2) + \log(|p^2 + m^2|/m^2) - 1 \\
& + (m^2/p^2) \log(|p^2 + m^2|/m^2)\}] \quad (5.9)
\end{aligned}$$

where

$$\begin{aligned}
B_1 &= \frac{1}{3}(\eta + 11)\beta_\mu^2 - 2 - \frac{1}{3}\eta, & B_2 &= -\frac{2}{3}\beta_\mu^2 - \frac{2}{3}(\eta + 2), \\
B_3 &= (\eta + 6)\beta_\mu^2 - \eta - 2, & B_4 &= -\eta - \frac{9}{2}, \\
B_5 &= -2(\eta + 7)\beta_\mu^2 + 2, & \eta &= -4 \cdot I_5 \quad \text{for scalar meson,} \\
& & &= -6 \cdot I_{10} \quad \text{vector meson.}
\end{aligned}$$

For the case $p^2 < -m^2$, the following term is added

$$\begin{aligned}
& (i\kappa^2/16\pi) \varepsilon(p_0) [(1/m) (p^2 B_1 - (\beta p)^2 B_2) (1 + (m^2/p^2)^3) \\
& + \{i(\beta p) B_4 - (1/m) (p^2 - m^2) B_5\} (1 - (m^2/p^2)^2) + m(1 + m^2/p^2) B_5]. \quad (5.9')
\end{aligned}$$

By the same procedure as the case of electron, the uncorrected propagation function of Boson $S_F^B(p) = \{- (1/m) (p^2 + m^2) + i(\beta p) + (1/m) (\beta p)^2\} / (p^2 + m^2)$ is replaced by the corrected one $S_F^{B'}(p)$ which behaves for $p^2 \gg m^2$

$$S_F^{B'}(p) \simeq 1/p^2. \quad (5.10)$$

This shows that for $p^2 \gg m^2$ the corrected propagation function damps faster than the uncorrected one by $1/p^2$. Thus, by using the corrected propagation functions of Fermion, photon and meson we can obtain the non-singular theory. Accordingly, there appears no divergence also in the case of the pseudo-scalar meson interacting with the nucleon field via the derivative coupling which is the interaction of the second kind.

§ 6. Concluding remarks

We have shown that the Hamiltonian for the entire system of gravitational field and matter field can be separated by eliminating the longitudinal and scalar graviton, in the following form

$$H = H_m + H_t + H_I + V$$

where H_m is the energy of matter field, H_t the energy of the transverse graviton, H_I the interaction term between the transverse graviton and matter field, and V the gravitational potential between the matters. For the macroscopical phenomena, only the potential

V , obtained by the unquantizable part of gravitational field, plays a role and by using this potential the equation of Einstein, Infeld and Hoffmann is deduced. That is, for the macroscopical phenomena, our quantization method of gravitational field is not inconsistent with the general relativistic treatments. On the other hand, the transverse graviton has no effect on the macroscopical phenomena. But, if we replace the free propagation function of matter field by the one corrected by the transverse graviton, it was useful to suppress the higher momentum states than $p \sim 1/\kappa (\sim 9 \times 10^{31} \text{ cm}^{-1} \sim 3 \times 10^{21} m_e, m_e: \text{electron mass})$, as seen from (5.6) and (5.7). And also, (4.22) or (4.29) shows that, at high momentum state $p \sim 1/\sqrt{G} \sim 2 \times 10^{32} m_e$, the gravitational potential becomes to be comparable with Coulomb potential. Therefore, it seems to us that the gravitational field will play an important role in the region where the internal structure of elementary particles comes into question.

The author wishes to express his sincere thanks to Prof. Y. Mimura for his kind interest in this work.

Appendix

Following Harish-Chandra,⁽¹⁾ we introduce such one-row vector matrix Γ_μ^* that $\Gamma_\mu^* \phi$ transforms as a vector for any Lorentz transformation. We now restrict our consideration to the case of neutral scalar particles. Then, there are following relations

$$\phi^\dagger \Upsilon_\mu = \Gamma_\mu^* \phi, \quad \bar{\phi} \partial_\mu \Upsilon_\nu = -\Gamma_\nu^* \beta_\mu \phi$$

$$\text{and} \quad \sqrt{2} \Gamma_\mu^* \phi = i/\sqrt{m} \cdot \partial_\mu V, \quad \sqrt{2} \Gamma_\mu^* \beta_\nu \phi = -i \partial_{\mu\nu} \sqrt{m} V$$

where $\sqrt{2}$ is due to the fact that, in the neutral field, only one canonical variable ϕ should be used, instead of both ϕ and $\bar{\phi}$.

Further, introduce a matrix β defined by

$$i\beta = \Upsilon_\mu \Gamma_\mu^*$$

Γ_μ^* , Υ_μ and β satisfy the following equations:

$$\Gamma_\mu^* \Upsilon_\nu = \delta_{\mu\nu}, \quad \Gamma_\mu^* \beta = \Gamma_\mu^*, \quad \beta^\dagger \Upsilon_\mu = \Upsilon_\mu, \quad \Gamma_\mu^* \beta_\lambda \Upsilon_\nu = 0$$

$$\beta_\mu \beta_\nu \Upsilon_\lambda = \Upsilon_\mu \cdot \delta_{\nu\lambda}, \quad \Gamma_\mu^* \beta_\nu \beta_\lambda = \delta_{\mu\nu} \Gamma_\lambda^*, \quad \beta \beta_\mu + \beta_\mu \beta = \beta_\mu; \quad \beta_\mu \beta \beta_\nu = \delta_{\mu\nu} (1 - \beta).$$

With the aid of the above relations, we can show that

$$\begin{aligned} \bar{\phi}(\beta h)[1 + (\beta n)^2](\beta h)\phi &= \bar{\phi}(\beta h)\beta[1 + (\beta n)^2](\beta h)\phi \\ &= \bar{\phi}(\beta h)\Upsilon_\lambda \Gamma_\lambda^*(\beta h)\phi + \bar{\phi}(\beta h)(n^\dagger \Upsilon)(n \Gamma^*)(\beta h)\phi \\ &= -(1/2m)[\{h_{\mu\lambda} - \frac{1}{2}\delta_{\mu\lambda}h\}\partial_\lambda V\}^2 + \{n_\mu(h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h)\partial_\nu V\}^2]. \end{aligned}$$

Similarly,

$$\begin{aligned} \bar{\phi}(\beta h)(\beta h)\phi &= -1/2m[\{h_{\mu\lambda} - \frac{1}{2}\delta_{\mu\lambda}h\}\partial_\lambda V\}^2 - (m^2/4)h^2V^2] \\ \bar{\phi}(\beta h)(\beta n)^2(\beta h)\phi &= -(1/2m)[\{n_\mu(h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h)\partial_\nu V\}^2 + (m^2/4)h^2V^2] \\ \bar{\phi}(\beta_\mu \beta_\nu - \frac{1}{2}\delta_{\mu\nu})\phi &= -(1/2m)[\partial_\mu V \cdot \partial_\nu V - \frac{1}{2}\delta_{\mu\nu}\{(\partial_\lambda V)^2 + m^2V^2\}]. \end{aligned}$$

References

- 1) A. Einstein, L. Infeld and B. Hoffmann, *Ann. Math.* **39** (1938), 65, Eq. (17•2). Cf. also B. Bertotti, *Nuo. Cim.* **12** (1954), 226.
- 2) De Donder, *La gravifique Einsteininne.* (Paris: Ganthiers-Villars) 1921.
- 3) A. Papapetrou, *Proc. Roy. Irish. Acad. A* **52** (1948), 11.
- 4) S. N. Gupta, *Phys. Rev.* **96** (1954), 1683.
- 5) S. N. Gupta, *Proc. Phys. Soc.* **65** (1952), 161, 608.
- 6) J. Belinfante, *Phys. Rev.* **98** (1955), 793.
- 7) S. Sakata, H. Umezawa and S. Kamefuchi, *Prog. Theo. Phys.* **7** (1952), 377.
- 8) C. N. Yang and D. Feldman, *Phys. Rev.* **79** (1950), 972.
- 9) H. Umezawa and Y. Takahashi, *Prog. Theo. Phys.* **9** (1953), 14, 501.
Y. Katayama, *Prog. Theo. Phys.* **10** (1953), 31.
- 10) Harish-Chandra, *Proc. Roy. Soc.* **186** (1946), 502.
- 11) Y. Nambu, *Prog. Theo. Phys.* **5** (1950), 614.
- 12) It has been shown also in the general relativity by Scheidegger, *Phys. Rev.* **82** (1951), 883.
- 13) N. Hu, *Phys. Rev.* **80** (1950), 1109.
- 14) S. Kamefuchi and H. Umezawa, *Prog. Theo. Phys.* **9** (1953), 529.
- 15) R. Arnowitt and S. Deser, *Phys. Rev.* **100** (1955), 349.
L. N. Cooper, *Phys. Rev.* **100**₂(1955), 362.

Note on the Decay Interactions of Hyperons and Heavy Mesons*

Chikashi ISO and Masaaki KAWAGUCHI

Physics Department, Tokyo University of Education, Tokyo

and

Research Institute for Fundamental Physics, Kyoto University, Kyoto

(Received May 12, 1956)

The decay processes are investigated based on Nishijima and Gell-Mann's charge independence theory of heavy unstable particles. We assume that the decay interaction Hamiltonian is a *spinor* in isotopic spin space, when all particles participating in a decay process have definite isotopic spins. The branching ratio of various decay modes including radiative decays and the relation of lifetimes among charge multiplets are obtained. For instance, the difference between lifetimes of θ^0 and of θ^+ , and the branching ratio of the τ^+ particle are well interpreted under our assumption. Our theory is also applied to the decay of the lightest hyperfragment ${}^3\text{H}_\Lambda$.

§ 1. Introduction

The recent developement of the investigations on heavy unstable particles from both experimental and theoretical sides presents one of the most interesting problems in physics. The curious natures of these particles are qualitatively well understood on the basis of the charge independence hypothesis proposed by Nishijima^{(2),(3),(4)} and Gell-Mann⁽⁵⁾. As far as the strong interactions are concerned their theory has surprisingly succeeded in explaining the experimental results.

In such a theory all elementary interactions are classified into three categories according to their strengths, that is,

a) Charge independent interactions

These are the strongest interactions for which the isotopic spin I as well as the η -charge (strangeness) are conserved, namely, $\Delta I = I^f - I^i = 0$, $\Delta \eta = \eta^f - \eta^i = 0$, where the superscripts i and f stand for the initial and final states, respectively.

b) Electromagnetic interactions

For these interactions, which are the strongest among those violating the charge independence, we have the selection rules $\Delta I = \pm 1, 0$, whereas $\Delta \eta = 0$.

c) Weak decay interactions

For weak decay interactions neither the isotopic spin I nor the η -charge are conserved. When all particles concerned have isotopic spins, the η -charge obeys the selection rule

* A preliminary report of the present paper has been published in this journal. See reference 1).

$\Delta J = \pm 1$, which was necessarily introduced in order to explain the cascade decay of a Ξ^- particle.

It is the aim of this paper to investigate the decay of heavy unstable particles in more detail on the basis of Nishijima and Gell-Mann's theory and of the following classification of the decay interactions.

Nishijima and one of the authors (M.K.) have previously proposed a further classification of the decay interactions into two classes¹⁾:

c') *Baryon decay interactions**

When all particles participating in a decay process have definite isotopic spins, the selection rule $\Delta J = \pm 1$ is equivalent to $\Delta I_3 = \pm 1/2$. In such a case we may assume that the decay interaction Hamiltonians are spinors in isotopic spin space. If we take this hypothesis for granted, the selection rules governing such decays are given by $\Delta I = \pm 1/2$, $\Delta I_3 = \pm 1/2$, which have first been proposed by Gell-Mann and Pais²⁾. We further assume that these interactions lead to lifetimes of the order of 10^{-10} sec. for two body decays.

c'') *Weaker decay interactions*

It is clear that not all decays are governed by the above mentioned interactions. For example, decays involving photons or leptons cannot be covered. Hence there must be weaker interactions which give lifetimes of the order of 10^{-8} sec. or longer. These interactions are, however, not always considered to be elementary interactions. Indeed, some of them are well explained in terms of the small deviation from interactions c'), for example, radiative corrections including the mass difference among charge multiplets.

Our classification is quite natural in the sense that weaker interactions cannot be composed of stronger interactions because of the difference of the transformation properties in isotopic space. In fact, the interactions mentioned above have the following properties in isotopic space,

- a) scalar,
- b) scalar + 3rd component of vector,
- c') one component of spinor.

The transformation property of the S-matrix for respective processes in isotopic spin space is also the same as that of the Hamiltonian, if one takes the lowest order approximation with respect to the interactions b) and c'). Therefore the discussion on the Hamiltonian c') can be regarded as that on the S-matrix of decay processes.

In § 2 we shall show some numerical results calculated from the interactions c'). Even if c') is firmly established, these results deviate inevitably from the experiments within 10% because of the influence of the interactions c''). Some of the curious natures of unstable particles, for example, the difference between lifetimes of θ^0 and of θ^+ , the branching ratio of the decay of τ^+ mesons, are well interpreted by the interactions c'). On the other hand, the relation between lifetimes of Σ^+ and of Σ^- would

* We have called c') as "charge independent decay interaction" in reference 1).

be one of the vital points in our theory when it is precisely measured in near future.

§ 3 is devoted to discuss the radiative decay processes of heavy particles. The radiative decays of unstable particles, except the case $\Delta I = \pm 1, 0$, do not occur by the interactions b), but they occur by the association of c') and b). If we consider, for instance, $\Sigma^+ \rightarrow p + \gamma$ in terms of c') and b), the branching ratio to other decay process is estimated from the experiments of photo-pion production and the general features of the S-matrix. The branching ratio thus obtained is different according to the type of the Σ particle. This result suggests us a possible method to determine the type of heavy unstable particles.

Further, our theory is applied to the hyperfragment ${}^3\text{H}_\Lambda$ in § 4. We find there the π^+ decay of ${}^3\text{H}_\Lambda$ occurs in a rate much smaller than $1/6$ when $I({}^3\text{H}_\Lambda) = 0$.

Detailed discussions on the Hamiltonian c') are given in Appendix.

Finally Nishijima and Gell-Mann's assignment of the isotopic spin and the η -charge for heavy particles is given in Table 1 for the sake of convenience.

Table 1

| Hyperons | I | η | Heavy Mesons | I | η |
|-----------|---------------|--------|----------------|---------------|--------|
| N | $\frac{1}{2}$ | 0 | π | 1 | 0 |
| Λ | 0 | -1 | θ, τ | $\frac{1}{2}$ | 1 |
| Σ | 1 | -1 | | | |
| Ξ | $\frac{1}{2}$ | -2 | | | |

§ 2. The baryon decay interactions

This section is devoted to discuss the *baryon decay interactions* in more detail. As is proposed in § 1, the interaction Hamiltonian is a spinor quantity in isotopic spin space, which gives the selection rules $\Delta I = \pm 1/2$, $\Delta I_3 = \pm 1/2$ and $\Delta \eta = \pm 1$. More rigorously, only one component of a spinor is adequate to the Hamiltonian here, because the other one breaks down the conservation of the electric charge. In what follows we shall exhibit the results derived from this hypothesis. Throughout this paper we do not enter into the spatial dependence of interactions.

i) $\Lambda \rightarrow N + \pi$

Under our assumption c') the interaction Hamiltonian must have the following form :

$$H \propto \sum_i \bar{\psi}(N) \varphi_i(\pi) \cdot \tau_i \varphi(\Lambda). \quad (2.1)$$

ψ denotes the wave function of a particle with isotopic spin $1/2$, and (N) , etc., serve to distinguish the particles concerned, while φ is the wave function of a particle with isotopic spin 0 or 1. τ 's are the current 2-2 isotopic spin matrices. Detailed discussions on the decay interaction Hamiltonian are given in Appendix. The Hamiltonian (2.1) gives the final system with $I' = 1/2$ only, and the branching ratio is^{1), 8)}

$$\frac{(\Lambda^0 \rightarrow p + \pi^-)}{(\Lambda^0 \rightarrow n + \pi^0)} = 2, \quad (2.2)$$

where $(\Lambda^0 \rightarrow p + \pi^-)$ means the number of events finding $p + \pi^-$.

Incidentally, if we take away the selection rule $\Delta I = \pm 1/2$, the branching ratio is given by^{8), 9)}

$$\frac{(\Lambda^0 \rightarrow p + \pi^-)}{(\Lambda^0 \rightarrow n + \pi^0)} = \frac{1 + 2x^2 + 2\sqrt{2}\rho x}{2 + x^2 - 2\sqrt{2}\rho x} \quad (2.3)$$

by introducing parameters

$$x = \pm |R_1|/|R_3| \quad \text{and} \quad \rho = \cos(\delta_3 - \delta_1), \quad (2.4)$$

where $|R_i|$ is hereafter the absolute value of the transition matrix element into the state $I' = i/2$, ($i=1$ and 3) and δ_i is the phase shift of pion-nucleon scattering in the isotopic spin state $i/2$. We can obtain the upper and lower limit of the branching ratio¹, but the range is too wide to mention particularly.

ii) $\Sigma \rightarrow N + \pi$

The transition matrix elements for the respective processes are similarly obtained as follows (see Appendix)

$$\langle p\pi^0 | R | \Sigma^+ \rangle = \sqrt{2/3} R_3 - \sqrt{1/3} R_1, \quad (2.5a)$$

$$\langle n\pi^+ | R | \Sigma^+ \rangle = \sqrt{1/3} R_3 + \sqrt{2/3} R_1, \quad (2.5b)$$

$$\langle n\pi^- | R | \Sigma^- \rangle = \sqrt{3} R_3. \quad (2.5c)$$

It is noted that the transition matrix elements of these three processes are described by two parameters R_1 and R_3 by virtue of our hypothesis.

The lifetimes of Σ^+ and Σ^- are

$$1/\tau(\Sigma^+) = R_3^2 + R_1^2, \quad (2.6a)$$

$$1/\tau(\Sigma^-) = 3R_3^2, \quad (2.6b)$$

or
$$\tau(\Sigma^-)/\tau(\Sigma^+) = (1 + x^2)/3. \quad (2.7)$$

In any case we get the inequality for lifetimes, namely,

$$3\tau(\Sigma^-) > \tau(\Sigma^+). \quad (2.8)$$

This is a criterion for the hypothesis of baryon decay interactions.¹

The branching ratio for the decays of Σ^{1+} is of the form of^{(7), (8), (9)}

$$\frac{(\Sigma^{1+} \rightarrow p + \pi^0)}{(\Sigma^{1+} \rightarrow n + \pi^+)} = \frac{2 + x^2 - 2\sqrt{2}\rho x}{1 + 2x^2 + 2\sqrt{2}\rho x}, \quad (2.9)$$

which is always valid even if one removes the assumption c'). From (2.9) one can easily see the following restrictions for the branching ratio,

$$\frac{1}{10} < \frac{(\Sigma^{1+} \rightarrow p + \pi^0)}{(\Sigma^{1+} \rightarrow n + \pi^+)} < 10, \quad \text{for } (3/2+) \quad (2.10a)$$

$$\frac{1}{100} < \frac{(\Sigma^{1+} \rightarrow p + \pi^0)}{(\Sigma^{1+} \rightarrow n + \pi^+)} < 100, \quad \text{for } (1/2+) \quad (2.10b)$$

$$\frac{1}{25} < \frac{(\Sigma^{1+} \rightarrow p + \pi^0)}{(\Sigma^{1+} \rightarrow n + \pi^+)} < 25, \quad \text{for } (1/2-) \quad (2.10c)$$

where $(3/2+)$ means the case that the Σ has spin $3/2$ and even parity. The values of the phase shifts of pion-nucleon scattering are taken as¹⁰⁾.

$$\begin{aligned} \delta_{33} &= 38.1^\circ, \quad \delta_{13} = 2.3^\circ, \quad \delta_{31} = 5.4^\circ, \quad \delta_{11} = -4.4^\circ, \\ \delta_8 &= -12.5^\circ \quad \text{and} \quad \delta_1 = 11.0^\circ. \quad (\sim 133 \text{ Mev.}) \end{aligned} \quad (2 \cdot 11)$$

Further, if we take the hypothesis of baryon decay interactions for granted, we can propose a method to determine the type of Σ particle⁷⁾. If we obtain accurate values of the branching ratio of the decay of Σ^+ particles and $\tau(\Sigma^-)/\tau(\Sigma^+)$ from experiments, the spin and the parity of Σ particle might be determined by (2.7) and (2.9). For a probable range of the branching ratio

$$(\Sigma^+ \rightarrow p + \pi^0)/(\Sigma^+ \rightarrow n + \pi^+) \approx 1,$$

the ratio of lifetimes is

$$\begin{aligned} \tau(\Sigma^-)/\tau(\Sigma^+) &\approx 1/3 \quad \text{or} \\ &5 \sim 10. \end{aligned} \quad (2 \cdot 12)$$

See, Fig. 1.

iii) $\Xi \rightarrow \Lambda + \pi$

In this case we only mention a result,

$$\tau(\Xi^0)/\tau(\Xi^-) = 2. \quad (2 \cdot 13)$$

iv) $\theta \rightarrow \pi + \pi$

The discussions differ essentially according to whether the θ particle has even spin and even parity or odd spin and odd parity.

a) *Even spin and even parity*

In this case the isotopic spin of the final state I^f must be 0 or 2 from Bose statistics, the latter case is forbidden from the selection rule $MI = \pm 1/2$. Then our baryon decay interaction forbids the decay $\theta^+ \rightarrow \pi^+ + \pi^0$. The experimental lifetimes of θ^+

$$\tau(\theta^+) \geq 10^{-8} \text{ sec.} \quad (2 \cdot 14)$$

is well understood under our assumptions. A possible interpretation of the decay of θ^+ in terms of the deviation from charge independence, including pion mass difference, will be given elsewhere.

On the other hand the decays of the θ^0 particle, which is allowed to take $I^f = 0$, have the branching ratio

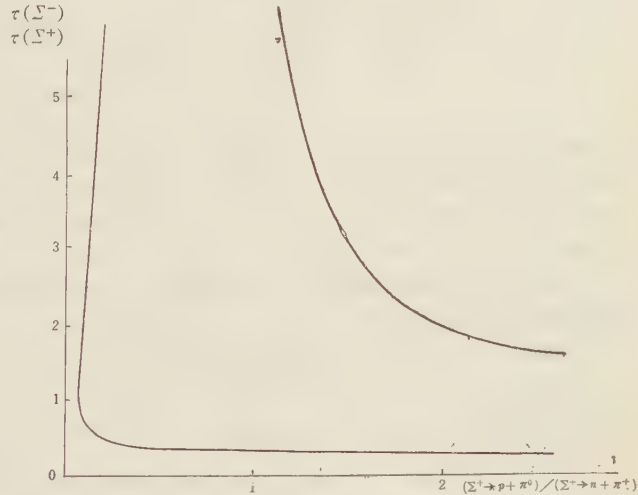


Fig. 1. Fig. 1 shows the relation between $(\Sigma^+ \rightarrow p + \pi^0)/(\Sigma^+ \rightarrow n + \pi^+)$ and $\tau(\Sigma^-)/\tau(\Sigma^+)$ for $(3/2+)$. For the cases $(1/2+)$ and $(1/2-)$ the curves almost coincide with that of $(3/2+)$.

$$(\theta^0 \rightarrow \pi^+ + \pi^-) / (\theta^0 \rightarrow \pi^0 + \pi^0) = 2. \quad (2 \cdot 15)$$

If we put the selection rule $\Delta I = \pm 1/2$ away, $I^f = 0$ as well as $I^f = 2$ are allowed. Then the branching ratio $(\theta^0 \rightarrow \pi^+ + \pi^-) / (\theta^0 \rightarrow \pi^0 + \pi^0)$ are expressed in terms of the phase shifts of pion-pion scattering like $(2 \cdot 3)$ or $(2 \cdot 9)^*$, and $\tau(\theta^+)/\tau(\theta_1^0) \geq 2$. Here we use Gell-Mann and Pais' result¹¹⁾ that we have two kinds of lifetimes of θ^0 and the shorter one of which is $\tau(\theta_1^0)$ here.

b) *Odd spin and odd parity*

In this case, only $I^f = 1$ is allowed and

$$\tau(\theta_1^0)/\tau(\theta^+) = 1.^{11)} \quad (2 \cdot 16)$$

In this case our hypothesis is, therefore, inadequate to explain the experimental results. Therefore θ must have even spin and even parity if our assumption is justified. It is needless to say that θ^0 cannot decay into two neutral pions in this case.

v) $\tau \rightarrow \pi + \pi + \pi$

Our assumption allows us to take only $I^f = 1$, and the branching ratio for τ^+ is given by^{12), 13)}

$$\frac{1}{4} \leq \frac{(\tau^+ \rightarrow 2\pi^0 + \pi^+)}{(\tau^+ \rightarrow 2\pi^+ + \pi^-)} \leq 1. \quad (2 \cdot 17)$$

This result agrees with the experiments.

The branching ratio has ever been calculated by Dalitz¹²⁾ under the assumption $I^f = I^f = 1$, and our result agrees with his one in spite of the different choice of I^f .

The relation between lifetimes of τ^+ and τ^0 is¹¹⁾,

$$\tau(\tau^+) \geq \tau(\tau_1^0), \quad (2 \cdot 18)$$

where $\tau(\tau_1^0)$ means the shorter one of two lifetimes of τ^0 . This relation is, however, valid only when τ has not any competing decay modes such as $\tau \rightarrow \mu + \nu$. If we neglect the state $I^f = 0$ for τ^0 , (2.18) becomes $\tau(\tau^+) = \tau(\tau_1^0)$.

§ 3. Radiative decay processes

In this section we shall study the radiative processes, which are concerned with the interaction c' , from the general features of the S-matrix and from well established experimental evidences. We do not have enough experimental data to observe the radiative decays of heavy unstable particles yet. In near future, however, a process like $\Sigma^+ \rightarrow p + \gamma$ will be observed and play an important role in investigating the detailed natures of heavy particles.

Nishijima and one of the authors (M.K.) have reported a method to know the type of the Σ particles⁷⁾. The general principle of the wave mechanics used is the unitarity of the S-matrix, which provides us the following relation,

$$\begin{aligned} & \langle p\gamma | R | p\pi^0 \rangle \langle p\pi^0 | R^* | \Sigma^+ \rangle + \langle p\gamma | R | n\pi^+ \rangle \langle n\pi^+ | R^* | \Sigma^+ \rangle \\ & = -2\text{Re} \langle p\gamma | R | \Sigma^+ \rangle. \end{aligned} \quad (3 \cdot 1)$$

We assume here that the radiative decay occurs at much smaller rate than the decay into a pion and a nucleon. We have previously proved that the matrix element of photo-pion production is expressed by its amplitude and the phase shifts of pion-nucleon scattering for respective isotopic spin states of the final system^(9,11). The results are given in angular momentum representation,

$$\langle p\pi^0 | R | p\gamma \rangle = \sqrt{2/3} |M_3| e^{i\delta_3} \mp \sqrt{1/3} |M_1| e^{i\delta_1}, \quad (3.2a)$$

$$\langle n\pi^+ | R | p\gamma \rangle = \sqrt{1/3} |M_3| e^{i\delta_3} \pm \sqrt{2/3} |M_1| e^{i\delta_1}, \quad (3.2b)$$

where M_i is the transition matrix element for the final state $I^J=i/2$. Then using (3.1), (3.2) and (2.5), we get

$$\frac{|\text{Re}\langle p\gamma | R | \sum^+ \rangle|^2}{|\langle n\pi^+ | R | \sum^+ \rangle|^2} = \frac{3}{4} \frac{|M_3|^2 + 2|M_1| \cdot |M_3| x + |M_1|^2 x^2}{1 + 2\sqrt{2} \rho x + 2x^2}, \quad (3.3)$$

where x and ρ are defined by (2.4). M_1 and M_3 are calculated from the experiments of photo-pion production, though there are some inevitable ambiguities.

Our fundamental relation (3.3) is valid *mutatis mutandis* for the radiative decays of other particles, and it is useful in order to see the branching ratio of radiative processes. Of course the information obtained from (3.3) is not complete, but we may make only a rough estimate. For the sake of convenience we show the state of pion-nucleon and photon-nucleon systems in Table 2.

Table 2

| Type of Hyperon | State of emitted pion-nucleon system | State of emitted photon-nucleon system |
|-----------------|--------------------------------------|--|
| 1/2- | $s_{1/2}$ | $E1$ |
| 1/2+ | $p_{1/2}$ | $M1$ |
| 3/2- | $d_{3/2}$ | $E1, M2$ |
| 3/2+ | $p_{3/2}$ | $M1, E2$ |
| .. | .. | .. |
| .. | .. | .. |
| .. | .. | .. |

i) $\Lambda^0 \rightarrow n + \gamma$

This process is quite unimportant from the practical point of view. Therefore we only give a rough estimation. Similar to (3.3) one can easily see under our assumption c') that

$$\frac{|\text{Re}\langle n\gamma | R | \Lambda^0 \rangle|^2}{|\langle p\pi^- | R | \Lambda^0 \rangle|^2} = \frac{3}{8} |M_1|^2, \quad (3.4)$$

where $|M_1|^2$ is found to be at most 1/1000 from the photo-pion cross sections. The imaginary part of the matrix element, though it is not rigorously calculated in our framework, is predicted to be of the same order of magnitude as its real part. Then the branching ratio of finding radiative decay is of the order of 1/1000.

ii) $\Sigma^+ \rightarrow p + \gamma$

On this process we have already reported in detail, so that we shall only repeat that the branching ratio of the radiative decay would serve to determine the type of Σ when one had thousands of experimental data on this particle. From (3.3) we have the branching ratio,

$$\begin{aligned} (\Sigma^+ \rightarrow p + \gamma) / (\Sigma^+ \rightarrow n + \pi^+) &= \text{order of } 1/100 \quad \text{for } (3/2+), \\ &= \text{order of } 1/1000 \quad \text{for } (1/2-), \\ &= \text{order of } 1/10000 \quad \text{for } (1/2+). \end{aligned} \quad (3.5)$$

Note that there are serious ambiguities, but the method mentioned here is quite general and is also valid if we remove the assumption c').

iii) $\theta^+ \rightarrow \pi^+ + \gamma$

When θ is a scalar particle, this process cannot occur. For a vector θ we have

$$\frac{|\text{Re} \langle \gamma \pi^+ | R | \theta^+ \rangle|^2}{|\langle \pi^0 \pi^+ | R | \theta^+ \rangle|^2} = \frac{1}{4} |M|^2, \quad (3.6)$$

where M is the matrix element of the hypothetical process $\gamma + \pi^- \rightarrow \pi^+ + \pi^-$ in the state $I^J=1$ with magnetic dipole radiation. However, we cannot get any instructive conclusion at present.

§ 4. Decays of hyperfragments

It is well known that a hyperfragment is such an excited state of a nucleus in which one neutron is substituted by the Λ^0 particle. The existence of the lightest hyperfragment ${}^3\text{H}_\Lambda$ is firmly established. In this section we shall apply our baryon decay interactions to the lightest hyperfragment ${}^3\text{H}_\Lambda$. The argument is different according as $I({}^3\text{H}_\Lambda)=0$ or $I({}^3\text{H}_\Lambda)=1$. The experiments are favorable to the former choice at present, because the partner ${}^3\text{He}_\Lambda$ does not seem to be well established yet.

We have already observed the following decay processes,

$$\begin{aligned} {}^3\text{H}_\Lambda &\rightarrow {}^3\text{He} + \pi^- \\ &\rightarrow d + p + \pi^- \\ &\rightarrow p + p + n + \pi^-. \end{aligned} \quad (4.1)$$

If we take the hypothesis of baryon decay interactions for granted, the branching ratios and other relations of the decays are easily obtained. Possible decay processes of ${}^3\text{H}_\Lambda$ and ${}^3\text{He}_\Lambda$ taken into consideration are,

$${}^3\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^- \quad (4.2a)$$

$$\rightarrow {}^3\text{H} + \pi^0 \quad (4.2b)$$

$$\rightarrow d + p + \pi^- \quad (4.2c)$$

$$\rightarrow d + n + \pi^0 \quad (4.2d)$$

$$\rightarrow p + p + n + \pi^- \quad (4 \cdot 2e)$$

$$\rightarrow p + n + n + \pi^0 \quad (4 \cdot 2f)$$

$$\rightarrow n + n + n + \pi^+ \quad (4 \cdot 2g)$$

and

$${}^3\text{He}_\Lambda \rightarrow {}^3\text{He} + \pi^0 \quad (4 \cdot 3a)$$

$$\rightarrow {}^3\text{H} + \pi^+ \quad (4 \cdot 3b)$$

$$\rightarrow d + p + \pi^0 \quad (4 \cdot 3c)$$

$$\rightarrow d + n + \pi^+ \quad (4 \cdot 3d)$$

$$\rightarrow p + p + p + \pi^- \quad (4 \cdot 3e)$$

$$\rightarrow p + p + n + \pi^0 \quad (4 \cdot 3f)$$

$$\rightarrow p + n + n + \pi^+ \quad (4 \cdot 3g)$$

i) Case of $I({}^3\text{H}_\Lambda) = 0$

In this case the corresponding ${}^3\text{He}_\Lambda$ does not exist. Our hypothesis c') allows $I^f = 1/2$ only, from which the following result is obtained,

$$\frac{({}^3\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^-)}{({}^3\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^0)} = \frac{({}^3\text{H}_\Lambda \rightarrow d + p + \pi^-)}{({}^3\text{H}_\Lambda \rightarrow d + n + \pi^0)} = 2. \quad (4 \cdot 4)$$

For four body decays,

$$({}^3\text{H}_\Lambda \rightarrow p p n \pi^-) = \frac{2}{3}A^2 + \frac{2}{3}B^2 + \frac{1}{6}C^2, \quad (4 \cdot 5a)$$

$$({}^3\text{H}_\Lambda \rightarrow p n n \pi^0) = \frac{1}{3}A^2 + \frac{1}{3}B^2 + \frac{1}{3}C^2, \quad (4 \cdot 5b)$$

$$({}^3\text{H}_\Lambda \rightarrow n n n \pi^+) = \frac{1}{2}C^2, \quad (4 \cdot 5c)$$

where

$$A = \alpha R_{1/2}(1 \frac{1}{2} 0), \quad B = \alpha R_{1/2}(1 \frac{1}{2} 1) \quad \text{and} \quad C = \alpha R_{1/2}(1 \frac{3}{2} 1).$$

Here $R_{1/2}(1 \frac{1}{2} j_1)$ is the transition matrix element of the resultant isotopic spin $1/2$ state, which is composed of the pion isotopic spin 1 and the sum of three nucleon isotopic spins j_2 . j_1 is the sum of the first and second nucleon isotopic spins, that is, $j_1 = 1/2 + 1/2$, and we further compose j_2 as the sum of the third nucleon isotopic spin and j_1 , namely, $j_2 = 1/2 + j_1$ and $1/2 = 1 + j_2$. α is a common factor independent of isotopic spins.

We get from (4.4) and (4.5)

$$({}^3\text{H}_\Lambda \rightarrow \pi^-) + ({}^3\text{H}_\Lambda \rightarrow \pi^+) = 2 \times ({}^3\text{H}_\Lambda \rightarrow \pi^0) \quad (4 \cdot 6)$$

and

$$\frac{1}{2} < \frac{({}^3\text{H}_\Lambda \rightarrow \pi^0)}{({}^3\text{H}_\Lambda \rightarrow \pi^-)} = \frac{1}{2 - \frac{3}{2}C^2} \leq 2. \quad (4 \cdot 7a)$$

If we do not find the π^+ decay of ${}^3\text{H}_\Lambda$,

$$({}^3\text{H}_\Lambda \rightarrow \pi^0) / ({}^3\text{H}_\Lambda \rightarrow \pi^-) = 1/2. \quad (4 \cdot 7b)$$

Further we assume that the branching ratio of two, three and four body decays are nearly equal, which is supported by the fact that respective processes are nearly equally observed in experiments. Then

$$(\text{Rate of finding four body decays}) = A^2 + B^2 + C^2 \approx 1/3. \quad (4.8)$$

Then $C^2 \ll 1/3$,

and therefore one can see from (4.5c)

$$({}^3\text{H}_\Lambda \rightarrow \pi^+) \ll 1/6. \quad (4.9)$$

From (4.4), (4.5a) and (4.5c)

$$\frac{({}^3\text{H}_\Lambda \rightarrow \pi^+)}{({}^3\text{H}_\Lambda \rightarrow \pi^-)} = \frac{\frac{1}{2}C^2}{\frac{2}{3} - \frac{1}{2}C^2} \ll \frac{1}{3}, \quad (4.10)$$

and further

$$\frac{1}{2} \leq \frac{({}^3\text{H}_\Lambda \rightarrow \pi^0)}{({}^3\text{H}_\Lambda \rightarrow \pi^-)} = \frac{\frac{1}{3}}{\frac{2}{3} - \frac{1}{2}C^2} \leq \frac{2}{3}. \quad (4.11)$$

ii) Case of $I({}^3\text{H}_\Lambda) = 1$

In this case the corresponding ${}^3\text{He}_\Lambda$ does exist, whose isotopic spin is also 1 and $I^3 = 1/2$ and $3/2$. The relation between the lifetimes is

$$1/2 \leq \tau({}^3\text{H}_\Lambda) / \tau({}^3\text{He}_\Lambda) \leq 2. \quad (4.12)$$

This is easily seen from the fact that the two body decays of ${}^3\text{H}_\Lambda$ and ${}^3\text{He}_\Lambda$ are described by common parameters, indeed, the matrix elements are given by

$$\langle {}^3\text{He}\pi^- | R | {}^3\text{H}_\Lambda \rangle = \sqrt{2/3} R_3 - \sqrt{1/3} R_1, \quad (4.13a)$$

$$\langle {}^3\text{H}\pi^0 | R | {}^3\text{H}_\Lambda \rangle = \sqrt{4/3} R_3 + \sqrt{1/6} R_1, \quad (4.13b)$$

and

$$\langle {}^3\text{He}\pi^0 | R | {}^3\text{He}_\Lambda \rangle = \sqrt{2/3} R_3 - \sqrt{1/3} R_1, \quad (4.13c)$$

$$\langle {}^3\text{H}\pi^+ | R | {}^3\text{He}_\Lambda \rangle = \sqrt{1/3} R_3 + \sqrt{2/3} R_1, \quad (4.13d)$$

where R_1 and R_3 are similar ones that are defined in § 2. For other processes, three and four body decays, the matrix elements are immediately written down in a similar manner. We get further from (4.13)

$$\frac{1}{2} \leq \frac{({}^3\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^-)}{({}^3\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^-) + ({}^3\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^0)} \bigg/ \frac{({}^3\text{He}_\Lambda \rightarrow {}^3\text{He} + \pi^0)}{({}^3\text{He}_\Lambda \rightarrow {}^3\text{He} + \pi^0) + ({}^3\text{He}_\Lambda \rightarrow {}^3\text{H} + \pi^-)} \leq 2. \quad (4.14)$$

The same relation holds for three body decays.

iii) ${}^4\text{H}_\Lambda$ and ${}^4\text{He}_\Lambda$

It seems to be reasonable to assume $I({}^4\text{H}_\Lambda) = I({}^4\text{He}_\Lambda) = \frac{1}{2}$. Then similar to the above cases, we get

$$({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^0 + n) + ({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^- + p) \geq \frac{1}{3}({}^4\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^- + n), \quad (4 \cdot 15a)$$

$$({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^- + p) + ({}^4\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^- + n) \geq ({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^0 + n), \quad (4 \cdot 15b)$$

$$({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^0 + n) + ({}^4\text{H}_\Lambda \rightarrow {}^3\text{He} + \pi^- + n) \geq \frac{1}{3}({}^4\text{H}_\Lambda \rightarrow {}^3\text{H} + \pi^- + p), \quad (4 \cdot 15c)$$

$$({}^4\text{H}_\Lambda \rightarrow d + p + n + \pi^-) \geq ({}^4\text{H}_\Lambda \rightarrow d + n + n + \pi^0). \quad (4 \cdot 15d)$$

For ${}^4\text{He}_\Lambda$ we have more complicated relations.

In conclusion we would like to express our thanks to Professor S. Hayakawa, Dr. K. Nishijima and Mr. Masatomo Sato for their kind guidance and valuable discussions. One of us (M.K.) is indebted to the Yukawa Fellowship of Osaka University for financial aid.

Appendix

As stated in § 1, the decay interaction Hamiltonian is assumed to be one component of a spinor in isotopic spin space. The decay interaction is so weak that the exact transition matrix element for a decay process has the same transformation property in isotopic spin space. Therefore the following discussions on the Hamiltonian are also valid for the S-matrix.

For $\Lambda \rightarrow N + \pi$, the possible combination, in order to construct a spinor quantity using their wave functions and τ -matrix, is only one, namely,

$$H \propto \sum_i \bar{\psi}(N) \varphi_i(\pi) \cdot \tau_i \varphi(\Lambda). \quad (\text{A} \cdot 1)$$

For $\Sigma \rightarrow N + \pi$, we have two possibilities,

$$\sum_i \bar{\psi}(N) \varphi_i(\pi) \cdot \varphi_i(\Sigma), \quad (\text{A} \cdot 2a)$$

and

$$\sum_i \bar{\psi}(N) (\varphi(\pi) \times \tau)_i \cdot \varphi_i(\Sigma). \quad (\text{A} \cdot 2b)$$

The Hamiltonians for $I^J = 1/2$ and $I^J = 3/2$ are immediately written by combining (A·2a), and (A·2b),

$$H(I^J = 1/2) = \sum_i (1/\sqrt{3}) R_i (\bar{\psi}(N) \varphi_i(\pi) \cdot \varphi_i(\Sigma) - \bar{\psi}(N) (\varphi(\pi) \times \tau)_i \varphi_i(\Sigma)), \quad (\text{A} \cdot 3a)$$

$$H(I^J = 3/2) = \sum_i (1/\sqrt{6}) R_i (2\bar{\psi}(N) \varphi_i(\pi) \cdot \varphi_i(\Sigma) + \bar{\psi}(N) (\varphi(\pi) \times \tau)_i \varphi_i(\Sigma)), \quad (\text{A} \cdot 3b)$$

from which we can obtain the matrix elements (2·5). (A·3) have some analogy to the transition matrix elements for photo-pion production.¹⁶⁾

It is straightforward to write down the Hamiltonian and the S-matrix of other processes. Of course the Hamiltonian concerned must be constructed also for such virtual process as $(\sum \sum K)$.

The transition matrix element of, say, $\Sigma \rightarrow N + \pi$ contains not only the direct interaction (Fig. 2.(i)) but also the complicated processes with higher order corrections of the strong interaction a) (Fig. 2.(ii), (iii)). After such higher order corrections by a), (A.3) are still valid when they are regarded as the transition matrix element.

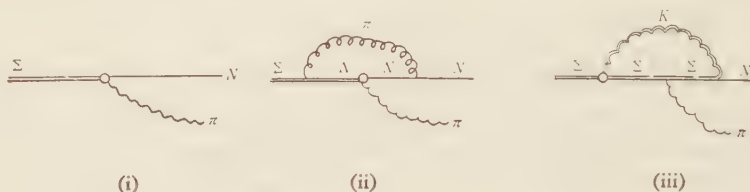


Fig. 2 Notation \bigcirc shows the weak baryon decay interactions.

References

- 1) M. Kawaguchi and K. Nishijima, *Prog. Theor. Phys.* **15** (1956), 180.
- 2) T. Nakano and K. Nishijima, *Prog. Theor. Phys.* **10** (1953), 581.
- 3) K. Nishijima, *Prog. Theor. Phys.* **12** (1954), 107.
- 4) K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285.
- 5) M. Gell-Mann, *Phys. Rev.* **92** (1953), 833.
- 6) M. Gell-Mann and A. Pais, *Proceedings of 1954 Glasgow Conference*, p. 342.
- 7) M. Kawaguchi and K. Nishijima, *Prog. Theor. Phys.* **15** (1956), 182.
- 8) G. Takeda, *Phys. Rev.* **101** (1956), 1547.
- 9) M. Kawaguchi and S. Minami, *Prog. Theor. Phys.* **12** (1954), 789.
- 10) F. de Hoffmann et al., *Phys. Rev.* **95** (1954), 1586.
- 11) M. Gell-Mann and A. Pais, *Phys. Rev.* **97** (1955), 1387.
- 12) R. H. Dalitz, *Proc. Phys. Soc.* **66A** (1953), 710.
- 13) G. Wentzel, *Phys. Rev.* **101** (1956), 1215.
- 14) K. M. Watson, *Phys. Rev.* **95** (1954), 228.
- 15) K. Brueckner and K. Watson, *Phys. Rev.* **86** (1952), 923.

Note added after the completion of the manuscript.

A similar theory has been developed by R. Gatto. (*Nuovo Cim.* **3** (1956) 318.)

Note added in proof.

By the recent experiments*) **) we have

$$\tau(\Sigma^-) \simeq (5 \sim 10) \times \tau(\Sigma^+) = 1.4^{+1.6}_{-0.5} \times 10^{-10} \text{ sec.}^*)$$

and

$$(\Sigma^{*+} \rightarrow p + \pi^0) / (\Sigma^{*+} \rightarrow n + \pi^+) \simeq 1.$$

These experimental values are in good agreement with the upper branch of the curve in Fig. 1. If we obtained more precise values of $\tau(\Sigma^-)/\tau(\Sigma^+)$ and $(\Sigma^{*+} \rightarrow p + \pi^0)/(\Sigma^{*+} \rightarrow n + \pi^+)$, it would be possible to determine the type of the Σ particles as stated in § 2.

*) W. D. Walker, 6th Rochester Conference. (1956).

**) W. F. Fry, 6th Rochester Conference. (1956).

W. F. Fry, J. Schneps, G. A. Snow and M. S. Swami, *Phys. Rev.* **100** (1955), 950.

On the Determination of the Spins of Λ^0 and θ^0 Particles

Shoichi HORI

Institute for Theoretical Physics, Kanazawa University, Kanazawa

(Received February 28, 1956)

A complete angular correlation function is derived, which is concerned with the directions of decay secondaries from Λ^0 and θ^0 particles produced in a π^-p collision and of the incident π^- -meson. The function is evaluated for a simplified case in which we assume that the spin of θ^0 is zero and the Λ^0 - θ^0 system is produced in S-state in center of mass system. For the case of the spin of $\Lambda^0=3/2$, the result agrees with that reported by L. Wolfenstein.

§ 1. Introduction

It is now well established¹⁾²⁾³⁾ that a Λ^0 and a θ^0 particle can be produced in a π^-p collision and they decay according to the schemes: $\Lambda^0 \rightarrow p + \pi^-$ and $\theta^0 \rightarrow \pi^+ + \pi^-$. Recently several attempts have been made to determine the spins of Λ^0 and θ^0 particles from the distribution of the angles between their production and decay planes.⁴⁾⁻¹¹⁾ In order to draw a clear-cut conclusion, however, further theoretical as well as experimental analyses seem to be necessary.

Before taking up the subject in question, we shall show the density matrices¹²⁾ of decay of Λ^0 and θ^0 particles* :

$$E_{dec}^{(\Lambda)}(m_s, m_s') = (-1)^{S-1/2} \tau^{-1} (2K+1) [(2S+1)/4\pi]^{1/2} \sum_{L_1 M_{L_1}} \langle K K O O | K K_1 L_1 O \rangle \cdot \langle S L_1 m_s M_{L_1} | S L_1 S m_s' \rangle W(KSKS; 1/2 L_1) Y_{L_1}^{M_{L_1}}(\mathbf{l}), \quad (1)$$

$$E_{dec}^{(\theta)}(m_s, m_s') = (2s+1)^{1/2} \tau'^{-1} (4\pi)^{-1/2} \sum_{L_2 M_{L_2}} (-1)^{s+L_2} \langle s s 0 0 | s s L_2 0 \rangle \cdot \langle s L_2 m_s M_{L_2} | s L_2 s m_s' \rangle Y_{L_2}^{M_{L_2}}(\mathbf{n}), \quad (2)$$

where

- S, m_s ; the spin of Λ^0 and its z-component,
- s, m_s ; the spin of θ^0 and its z-component,
- \mathbf{l} ; direction of relative momentum between decay secondaries of Λ^0 (π^- and p),
- \mathbf{n} ; direction of relative momentum between decay secondaries of θ^0 (π^+ and π^-),

* $\langle | \rangle$ and W here denote the Clebsh-Gordan and Racah coefficients respectively.

τ, τ' ; life times of A^0 and θ^0 .

In eq. (1), summation is performed with respect to the direction of the spin of the outgoing proton. K , the relative orbital angular momentum of the decay secondaries π^- and p , takes either one of the values $S \pm 1/2$ according to the parity of A^0 relative to that of a proton. It is obvious that $|M_{L_1}|_{max} = L_1$, $|M_{L_2}|_{max} = L_2$, while the values L_1 and L_2 are restricted by the Clebsch-Gordan and Racah coefficients in the regions: $0 \leq L_1 \leq 2S$ or $2S-1$, $0 \leq L_2 \leq 2s$, so that distributions of azimuthal angles of l and n , i.e., angles between the production plane and the decay planes will determine the spins of A^0 and θ^0 .

The following fact should however be noted. Let us now ignore the spins of all particles in consideration except for a decaying particle, say, θ^0 -particle and we shall calculate the density matrix of the production in the centre of mass system and take the direction of A^0 -particle as z -axis and let the direction of the incident π^- -meson be k . If we assume moreover that the A^0 - θ^0 system is produced in S -state, then the density matrix becomes the complex conjugate of eq. (2) apart from a constant factor and the angular correlation function w becomes¹²⁾

$$\begin{aligned} w(\mathbf{n}, \mathbf{k}) &= \sum_{m_s, m_s'} E_{dec}^{(0)}(m_s, m_s') E_{prod}(m_s, m_s') \\ &\cong |Y_s^0(\mathbf{n} \cdot \mathbf{k})|^2 \\ &\cong \sum_{L_2=0}^{2s} (2L_2+1)^{-1/2} [\langle s s 0 0 | s s L_2 0 \rangle]^2 Y_{L_2}^0(\mathbf{n} \cdot \mathbf{k}). \end{aligned} \quad (3)$$

(\cong means the equality holds apart from a constant factor.)

Therefore \mathbf{n} distributes in axially symmetric manner around \mathbf{k} and s can be determined from the distribution of the azimuthal angle $\cos^{-1} \mathbf{n} \cdot \mathbf{k}$. Of course if we average eq. (3) with respect to the direction (zenith angle) of \mathbf{k} , we shall obtain the distribution (2) again, but it is obvious that the method is at a disadvantage from the statistical point of view. Thus we can see from the simplified consideration that it is desirable to calculate the density matrix of the production and to obtain the complete angular correlation function of l , \mathbf{n} and \mathbf{k} . This will be done in the next section but as the numerical evaluation of the general formula is very tedious, we shall evaluate it only for a particularly simple case in Section 3. Comparison with experiments will not be done, because we could not find enough data of associated productions with hydrogen used as target, from which we could have deduced quantities necessary for the comparison.

§ 2. The density matrix of the production and the angular correlation function

Let the directions of relative angular momenta of the incident π^- -meson and proton and of the outgoing A^0 and θ^0 be \mathbf{k} and \mathbf{p} respectively in the centre of mass system

† Smaller one of $2K$ and $2S$.

and the z-component of the spin of the proton be μ . Then the matrix element of the transition is given by

$$\begin{aligned} (\mathcal{S}m_S; sm_s; \mathbf{P} | \tfrac{1}{2} \mu; \mathbf{k}) &= \sum_{lm} \sum_{l_1 m_1} a(J j l l_1) Y_l^{m*}(\mathbf{P}) \\ &\cdot (\mathcal{S}m_S; sm_s; lm | S(j) l J M_J) \\ &\cdot \langle \tfrac{1}{2} l_1 J M_J | \tfrac{1}{2} l_1 \mu m_1 \rangle Y_{l_1}^{m_1}(\mathbf{k}), \end{aligned} \quad (4)$$

where

$$\begin{aligned} l, m; & \text{ relative ang. mom. and its } z\text{-comp. of the } \Lambda^0 - \theta^0 \text{ system,} \\ l_1, m_1; & \text{ relative ang. mom. and its } z\text{-comp. of the } \pi^- - p \text{ system, and} \\ (\mathcal{S}m_S; sm_s; lm | S(j) l J M_J) &= \sum_{m_j} \langle S s m_S m_s | S s j m_j \rangle \\ &\times \langle j l m_j m | j l J M_J \rangle. \end{aligned}$$

$a(J j l l_1)$ is a coefficient which depends on the incident energy and the detailed properties of the interaction. In particular if we choose the direction of \mathbf{P} as z axis, we have

$$\begin{aligned} (\mathcal{S}m_S; sm_s; \mathbf{P} | \tfrac{1}{2} \mu; \mathbf{k}) &= \sum_l \sum_{l_1 m_1 j} \sum_{M_J} a(J j l l_1) [(2l+1)/4\pi]^{1/2} \\ &(\mathcal{S}m_S; sm_s; l 0 | S(j) l J M_J) \langle \tfrac{1}{2} l_1 J M_J | \tfrac{1}{2} l_1 \mu m_1 \rangle Y_{l_1}^{m_1}(\mathbf{k}). \end{aligned} \quad (5)$$

After averaging with respect to μ , the density matrix becomes as follows,

$$\begin{aligned} E_{prod}(m_S, m_S'; m_s, m_s'; \mathbf{k}) &= \tfrac{1}{2} \sum_{\mu} (\mathcal{S}m_S; sm_s; \mathbf{P} | \tfrac{1}{2} \mu; \mathbf{k}) \\ &\cdot (\mathcal{S}m_S'; sm_s'; \mathbf{P} | \tfrac{1}{2} \mu; \mathbf{k})^* = \tfrac{1}{2} \sum_l (-1)^{J'-1/2-l-l_1'} \\ &\cdot [(2l+1)(2l'+1)(2l_1+1)(2l_1'+1)(2J'+1)]^{1/2} a(J j l l_1) a^*(J' j' l' l_1') \\ &\cdot \langle l_1' l_1 0 0 | l_1' l_1 L 0 \rangle (\mathcal{S}m_S; sm_s; l 0 | S(j) l J M_J) \\ &\cdot (\mathcal{S}m_S'; sm_s'; l' 0 | S(j') l' J' M_J') \langle J' L M_J M_L | J' L J M_J \rangle \\ &\cdot W(J' l_1' J l_1; \tfrac{1}{2} L) Y_L^M(\mathbf{k}). \end{aligned} \quad (6)$$

The summation here is over $l, l', l_1, l_1', j, J, M_J, j', J', M_J'$, and LM . A relation between the total cross section σ and the coefficient $a(J j l l_1)$ can be derived from eq. (6);

$$\begin{aligned} \sigma &= (1/v) \int \sum_{m_S m_s} E_{prod}(m_S, m_S'; m_s, m_s'; \mathbf{k}) d\mathcal{Q}_k \\ &= (1/2v) (4\pi)^{1/2} \sum_{j J l l_1} (-1)^{l+l_1} (2J+1) |a(J j l l_1)|^2, \end{aligned} \quad (7)$$

where \mathcal{Q}_k is a solid angle along the direction of \mathbf{k} and v , the incident flux ($1/\text{cm}^2 \text{ sec}$).

As the angular correlation function is given by

$$\begin{aligned} w(\mathbf{l}, \mathbf{n}, \mathbf{k}) &= \sum_{m_S m_S'} \sum_{m_s m_s'} E_{dec}^{(\Lambda)}(m_S, m_S') E_{dec}^{(0)}(m_s, m_s') \\ &\cdot E_{prod}(m_S, m_S'; m_s, m_s'), \end{aligned}$$

if we insert eqs. (1), (2) and (6) into this equation and define

$$\begin{aligned} & [L_1 L_2 M_{L_1} M_{L_2} | S s j j' l l' | J M_J, J' M_{J'}] \\ &= \sum_{m_s m_s'} \sum_{m_s m_s'} \langle S L_1 m_s M_{L_1} | S L_1 S m_s' \rangle \langle s L_2 m_s M_{L_2} | s L_2 s m_s' \rangle \\ & \cdot (\delta m_s; m_s; l O | S s(j) l J M_J) (\delta m_s'; m_s'; l' O | S s(j') l' J M_{J'}), \end{aligned} \quad (8)$$

and

$$\begin{aligned} & \{L_1 L_2 M_{L_1} M_{L_2} | S s j j' l l' J J' | L_1 L_2 L - M_L\} \\ &= \sum_{M_J M_{J'}} [L_1 L_2 M_{L_1} M_{L_2} | S s j j' l l' | J M_J, J' M_{J'}] \\ & \cdot \langle J' L M_{J'} M_L | J' L J M_J \rangle, \end{aligned} \quad (9)$$

then we obtain

$$\begin{aligned} w(\mathbf{l}, \mathbf{n}, \mathbf{k}) &\cong \sum (-1)^{J-1/2-l'-l} [(2l+1)(2l'+1)(2l_1+1)(2l_1'+1)(2J+1)]^{1/2} \\ & \cdot a(J j l l_1) a^*(J' j' l' l'_1) \langle s s O O | s s L_2 O \rangle \langle K K O O | K K L_1 O \rangle W(K S K S; 1/2 L_1) \\ & \cdot Y_{L_1}^{M_{L_1}}(\mathbf{l}) Y_{L_2}^{M_{L_2}}(\mathbf{n}) \{L_1 L_2 M_{L_1} M_{L_2} | S s j j' l l' J J' | L_1 L_2 L - M_L\} \\ & \cdot \langle l'_1 l_1 O O | l'_1 l_1 L O \rangle W(J' l'_1 J l_1; 1/2 L) Y_L^{M_L}(\mathbf{k}). \end{aligned} \quad (10)$$

Here the summation is over $l l', l_1 l'_1, j J M_J, j' J' M_{J'}, L_1 M_{L_1}, L_2 M_{L_2}$ and LM . It must be noted here that eq. (9) has a nonvanishing value only for $M_{J_1} + M_{J_2} = -M_L$. It is very tedious to evaluate eq. (10) for the general case mainly because of interference of states with diverse angular momenta. It may be better to recalculate the density matrix $E_{pp,nt}$ in the lowest order of the perturbation theory.* In the next section we shall evaluate eq. (10) for a simplified case.

§ 3. A special case

Here we make the following assumptions,

- i) L^0 and θ^0 are produced with zero relative ang. mom. in the center of mass system.
- ii) the spin s of θ^0 is zero.

As L^0 and θ^0 are rather heavy, the assumption i) will be good, unless the interaction is of a particular type. And if the assumption ii) is valid, \mathbf{n} will show an isotropic distribution independently of the production mechanism. There is an experimental evidence which may support the assumption ii).¹⁰⁾

In this case the angular correlation function reduces to

$$\begin{aligned} w(\mathbf{l}, \mathbf{k}) &= (v\sigma/\tau) (S^2/\pi) \sum_{L,M} (2L+1)^{-1} \\ & \times [\langle K K O O | K K L O \rangle W(K S K S; 1/2 L)]^2 Y_L^{M*}(\mathbf{l}) Y_L^M(\mathbf{k}), \end{aligned} \quad (11)$$

* Of course it is dubious to use the perturbation approximation for such a strong interaction. It may however be granted, if we confine ourselves to the problem of angular correlation and to the vicinity of a threshold energy. Such a calculation will be made in a subsequent paper.

taking eq. (7) into account. Here we put $l_1=K=S-1/2$. The other choice will lead to the same conclusion, as we shall see more clearly in the next section. Using the addition theorem of the spherical harmonics, we find

$$\begin{aligned} (\tau/v\sigma)w(\mathbf{l}, \mathbf{k}) &= (S^2/\pi) \sum_L [4\pi(2L+1)]^{-1/2} Z_S^2(L) Y_L^0(\mathbf{l} \cdot \mathbf{k}) \\ &= (S/2\pi)^2 \sum_L Z_S^2(L) P_L(\cos \theta), \end{aligned} \quad (12)$$

where we put

$$Z_S(L) = \langle KKOO | KKLO \rangle W(KSKS; 1/2L), \quad (13)$$

and θ is the angle between \mathbf{l} and \mathbf{k} , (Fig. 1). The right hand side of Eq. (12) is a simple polynomial of $x=\cos \theta$. These are,

$$\begin{aligned} (32\pi^2\tau/v\sigma)w(\cos \theta) &= 1 && \text{for } S=1/2 \\ &= 1/4(3x^2+1) && \text{for } S=3/2 \\ &= 1/4(5x^4-2x^2+1) && \text{for } S=5/2 \end{aligned}$$

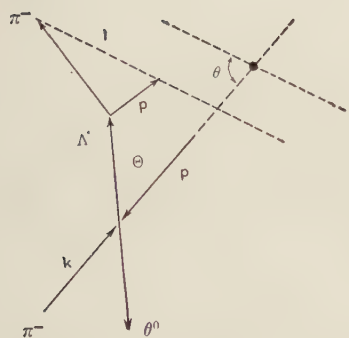


Fig. 1.

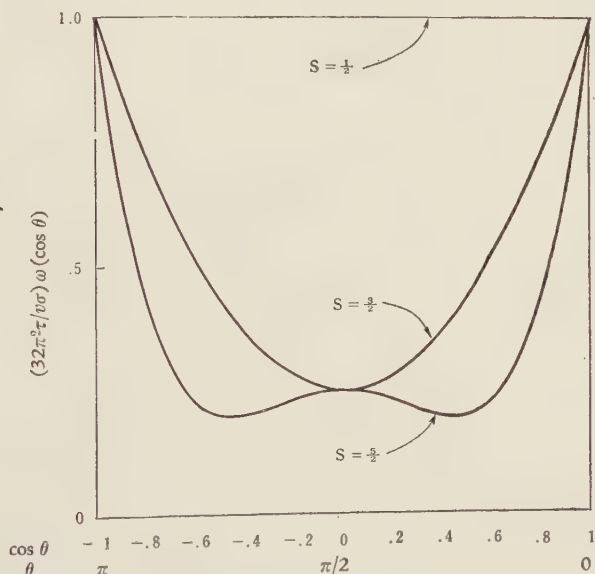


Fig. 2. Angular distribution of θ (Fig. 1) in the case in which the spin of θ^0 is zero and $\Lambda^0 - \theta^0$ system is produced in S -state in center of mass system.

The plot of these functions is shown in Fig. 2. The case for $S=3/2$ agrees with that reported by L. Wolfenstein.¹¹⁾ For the case $S>1/2$ we see that the decay plane inclines to coincide with the production plane except for a considerably small angle θ .

§ 4. Discussions

The result in the previous section can be derived from a simpler and rather intuitive

consideration.

Let θ be the angle between z -axis arbitrarily chosen and the direction of relative momentum l of the decay secondaries (π^- , p) of a J^0 -particle. When the z -component of the spin of J^0 takes a definite value m_s , the angular distributions $f_{m_s}(\theta)$ can be easily obtained by setting $m_s' = m_s$ in eq. (1) (apart from a numerical factor): these are (irrespective of the two values of $K = S \pm 1/2$)^{*}

$$\left. \begin{aligned} f_{\pm 1/2}(\theta) &= 1 && \text{for } S = \frac{1}{2} \\ f_{\pm 3/2}(\theta) &= \frac{3}{2}(1-x^2) \\ f_{\pm 1/2}(\theta) &= \frac{1}{2}(1+3x^2) \end{aligned} \right\} \text{for } S = \frac{3}{2}$$

$$\left. \begin{aligned} f_{\pm 5/2}(\theta) &= \frac{15}{8}(1-x^2)^2 \\ f_{\pm 3/2}(\theta) &= \frac{3}{8}(1-x^2)(1+15x^2) \\ f_{\pm 1/2}(\theta) &= \frac{3}{4}(1-2x^2+5x^4) \end{aligned} \right\} \text{for } S = \frac{5}{2}$$

etc.

where $x = \cos \theta$.

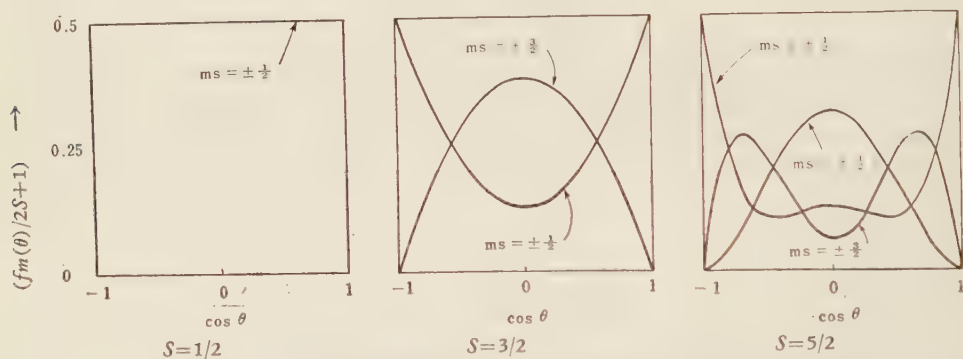


Fig. 3 Angular Distributions $f_{m_s}(\theta)$ (m_s should be read as m_s .)

We can afford a simple interpretation to these results; in the case of $S = \frac{1}{2}$, the spin S of J^0 is (anti)parallel to the z -axis when $m_s = \pm S$ and it is nearly perpendicular to the z -axis when $m_s = \pm \frac{1}{2}$, and the relative momentum l , which is perpendicular to the relative angular momentum of the π - p system, is apt to become perpendicular to S because of the conservation of angular momentum.

When the direction of S is completely unknown, we must average the angular distributions with respect to m_s , obtaining

$$(1/2S+1) \sum_{m_s} f_{m_s}(\theta) = 1.$$

Thus it becomes isotropic independently of S . Therefore we cannot determine S from

* The value of $(2K+1) \langle KKOO | KKLO \rangle W(KSKS; 1/2L)$ is the same for $K = S \pm 1/2$

the angular distribution of θ , if we confine ourselves only to decays of Λ^0 .

S. B. Treiman et al.⁴⁾ studied the distribution of the angle between decay plane of Λ^0 , assuming that the spin of a Λ^0 produced in a π^- -p collision is always perpendicular to the direction of propagation of the Λ^0 .^{*} Their assumption is not exactly correct and so are the angular distributions they obtained, although their conclusion that the two planes are inclined to coincide with each other in the case of $S > \frac{1}{2}$ will not be altered under a more exact treatment.

Let us suppose that the incident π^- -meson comes in along the z-axis and Λ^0 , θ^0 and secondaries (π^+ , π^-) of the θ^0 go out along the \pm z-axis. Then the value of m_s is equal to the z-component μ of the spin of the initial proton and is confined to $\pm \frac{1}{2}$, because the z-component of the total angular momentum should conserve. Therefore the angular distributions in this case agree with what we showed in the figure 2 (cf. Fig. 3),^{**} and S will be determined from these distributions, when more of available data are accumulated in future. This is the case treated by R. K. Adair. Since the data of associated productions to be adopted in this case are much restricted, the method seems to be at a disadvantage from the statistical point of view. However if we learn from other considerations that the spin of θ^0 is zero and Λ^0 and θ^0 are produced with vanishing relative angular momentum, there need not be such a restrictions. This is just the case we treated in § 3.

Though these assumptions i) and ii) in (§ 3) seem to be very plausible,¹³⁾¹⁴⁾ if this is not the case, we are obliged either to adopt the method suggested by R. K. Adair, or to compare experimental data with our more general formula (10) in (§ 2).

The author wishes to express his gratitude to Prof. T. Tati, not only for his criticism, but also for his interest and encouragement. The author is also indebted to Mr. A. Wakasa and Mr. T. Tani for their assistance in the working out of the problem.

References

- 1) W. B. Fowler et al., Phys. Rev. **91** (1953), 1287; **93** (1954), 861; **98** (1955), 121.
- 2) R. W. Thompson et al., Phys. Rev. **95** (1954), 1576.
- 3) W. D. Walker, Phys. Rev. **98** (1955), 1407.
- 4) S. B. Treiman et al., Phys. Rev. **97** (1955), 244; *ibid.* **100** (1955), 879.
- 5) J. Ballam et al., Phys. Rev. **97** (1955), 245.
- 6) G. D. James et al., Phil. Mag. **46** (1955), 571.
- 7) J. D. Sorrels, Proc. Fifth Annual Rochester Conf. (1955), 90.
- 8) P. T. Matthew and A. Salam, Phil. Mag. **46** (1955), 150.
- 9) R. K. Adair, Phys. Rev. **100** (1955), 1540.
- 10) W. D. Walker and W. D. Shephard, (unpublished). The author is indebted to Drs. W. D. Walker and W. D. Shephard who sent him their manuscript before publication.
- 11) L. Wolfenstein, Phys. Rev. **94** (1954), 150.
- 12) L. C. Biedenharn and M. E. Rose, Rev. Mod. Phys. **25** (1953), 729.
- 13) M. Kawaguchi and K. Nishijima, Prog. Theor. Phys. **15** (1956), 180.

* In reality this happens merely with a certain probability.

** Mixing of two states with $\mu = \pm 1/2$ produces no interference, since two amplitudes corresponding to $m_s = \pm 1/2$ are equal to each other.

C. Iso, *Soryushiron Kenkyu*, **11** (1956), 1.

14) G. Takeda, Private communication; *Soryushiron Kenkyu*, **11** (1956), 90.

Note added in proof

Recently S. B. Treiman (Phys. Rev. 101 (1956), 1216) suggested a method for determining spins of hyperons from the reaction $K^- + p \rightarrow \Lambda^0 + \pi^0$ (or $\Sigma^0 + \pi^-$). The method can be explained also from the same simple consideration we made in § 4. In this case we have only to choose the line of flight of the hyperon as z -axis.

Partial Description of Quantum-Mechanical System

Gentaro ARAKI

Faculty of Engineering, Kyoto University, Kyoto

(Received March 30, 1956)

A quantum-mechanical system is divided into several partial systems which are interacting each other. The general and practical methods of extracting one of the partial systems are discussed. If a one-dimensional partial system is extracted the two-particle interaction in the partial system is a long range part of the two-particle interaction in the original system and its short range part is absorbed in the average field due to the remaining part of the system. It is noted that the light absorption of long carotenoids is due to the plasma oscillation of π electrons caused by the long range part of the Coulomb interaction. The mechanism of screening the short range part is somewhat different from that of the three-dimensional case.

Introduction

As has been generally realized, it is hopelessly difficult to obtain the exact solution for most quantum-mechanical problems of real significance. Various ways for obtaining approximate solutions have been devised. Though they are only mathematical in some case, the device often introduces the method of physically important significance in which the approximation means a kind of the physical idealization. In such a case the physically significant conclusion can be drawn by the idealization and realization (the analysis and synthesis) which have been the general method in the science.

There are two ways for the idealization. The first is to introduce the physically ideal basic states. For example, one has often preferred the idea of configurations to the neat mathematical expression for studying nuclear, atomic and molecular states. The second is to divide a given quantum-mechanical system into several parts which are interacting each other, as is the case with molecules which have preferably been considered as consisting of atoms in order to fit the chemical idea of elements and compounds. The division can be done with respect to the coordinates as is the case with the use of relative and centre-of-mass coordinates in various problems.

In a certain case only one partial system is extracted and the state of the remaining parts is made asleep by an appropriate way. This may be called the method of the partial description. The sleeping partial system is often called a core or a skeleton. The energy levels of atoms are usually discussed by making the state of their nuclei sleep and the latter is ignored from the outset. This may be an extreme, trivial case, while the idea of valence electrons outside a closed atomic core may be a better example. The separation of the electronic and nuclear systems in molecular problems and extracting the π -

electron system from the large conjugated molecules are the same kind of examples. A special way of the partial description is to extract a system of partial coordinates from the partial or total system after the suitable coordinate transformation. The separation of vibrational and rotational coordinates of molecules and the free-electron model for the π -electron system^{1,2,3,4} belong to this category. The theory of collective motion^{5,6)} is also of the same kind.

In what follows a general method to extract the partial system will first be discussed in order to see what is the kind of its approximation. The method itself will not be a new one but a generalization of those used in various cases. The practical method will be accounted for in the cases of helium and complex atoms. It will next be shown that the one-dimensional partial description results in dividing the two-particle interaction into the short and long range parts and adopting the latter as the two-particle interaction in the extracted partial system, the short range part being absorbed in the skeleton field. Though this will be shown for the Coulomb interaction in the present paper, the result can be generalized to the case of other types of interactions. The theory of carotenoids¹⁾ and polyacenes³⁾ previously worked out by the author and his collaborator will be interpreted from this viewpoint. That this method is really useful will further be shown with actual examples in the separate papers in which we shall see that the result is in fairly good agreement with experiment in spite of the simpleness of the model and the calculation.

Finally it will be noted that the light absorption by long carotenoids is really caused by the plasma oscillation of π electrons contained in the conjugated chains. This plasma oscillation is excited by the long range part of the Coulomb interaction between π -electrons but the reason of screening is somewhat different from that of the three-dimensional case⁵⁾. It is due to the excitation of only a one-dimensional motion and sleeping of the motion in the remaining two dimensions, while the three-dimensional plasma oscillation is caused by the charge screening which results in a cut-off of the high frequency Fourier components of the Coulomb interaction. However, the one-dimensional and three-dimensional cases may both be considered as belonging to the same kind of phenomena from the viewpoint of the partial description.

The present idea of the partial description by extracting the partial two-particle interaction could successfully be applied to various problems, for example, including nuclear problems as well as atomic and molecular ones, if the method of extracting the partial system and the partial two-particle interaction would ingeniously be devised.

§ 1. General method

We consider a quantum-mechanical system of N coordinates. After an appropriate transformation of the coordinates we divide all coordinates into two sets: $(\xi_1, \xi_2, \xi_3, \dots, \xi_n)$ and $(x_1, x_2, x_3, \dots, x_{N-n})$. Their canonically conjugate momenta are denoted by $(\pi_1, \pi_2, \pi_3, \dots, \pi_n)$ and $(p_1, p_2, p_3, \dots, p_{N-n})$ respectively. For the sake of simplicity we shall denote these sets of coordinates and momenta by ξ, x, π , and p respectively. The Hamiltonian of the total system consequently takes the following form:

$$H = T(\pi, \xi) + V(\xi) + H_0(p, x) + H_1(\pi, \xi; p, x). \quad (1.1)$$

The total system is thus divided into two partial systems which are represented by the Hamiltonians denoted by $T(\pi, \xi) + V(\xi)$ and $H_0(p, x)$. The interaction between these two systems is given by $H_1(\pi, \xi; p, x)$.

In order to extract a partial system we have to make the other sleep. This process can be carried out in the following way. First we solve the eigenvalue problem:

$$[H_0(p, x) + H_1(0, \xi; p, x)]\phi(\xi, x) = U(\xi)\phi(\xi, x) \quad (1.2)$$

where ξ is considered as the constant parameter, $U(\xi)$ is the eigenvalue, and ϕ is normalized for every value of ξ . If a set of all $\phi(\xi, x)$ as the functions of x is complete for every value of ξ , any eigenfunction Ψ of the total Hamiltonian can be expanded in an infinite series of $\phi(\xi, x)$ for every value of ξ :

$$\Psi(\xi, x) = \sum_{\xi} \psi(\xi) \phi(\xi, x). \quad (1.3a)$$

The coefficient $\psi(\xi)$ of the expansion can be so determined that $(H - E)\Psi = 0$ is satisfied by the series. For the present purpose we should assume that the series consists of only a few terms. For the sake of simplicity we consider the simplest case in which the series consists of a single term:

$$\Psi(\xi, x) = \psi(\xi) \phi(\xi, x). \quad (1.3b)$$

Of course this can not be an exact solution but an approximate one for the eigenvalue problem of the total Hamiltonian, except for the accidentally fortunate case.

The eigenvalue equation, $(H - E)\Psi = 0$, for the total system reduces to

$$[T(\pi, \xi) + V(\xi) + U(\xi) + H_1(\pi, \xi; p, x) - H_1(0, \xi; p, x) - E]\psi(\xi) \phi(\xi, x) = 0. \quad (1.4)$$

Further if we multiply this by ϕ^* and integrate with respect to x , we have the equation for ψ as follows:

$$[T(\pi, \xi) + V(\xi) + U(\xi) + T_1(\pi, \xi) - E]\psi(\xi) = 0 \quad (1.5a)$$

where

$$\begin{aligned} T_1(\pi, \xi) = & \int \phi^*(\xi, x) [H_1(\pi, \xi; p, x) - H_1(0, \xi; p, x)] \phi(\xi, x) dx \\ & + \int \phi(\xi, x) [T(\pi, \xi) \phi(\xi, x) - \phi(\pi, \xi) T(\pi, \xi)] dx. \end{aligned} \quad (1.5b)$$

Generally speaking, $\psi(\xi)$ must contain x as a parameter in order to satisfy (1.4). This contradicts the starting assumption, (1.3b). Therefore the influence of T_1 on the eigenvalue problem should be very small in order that the assumption given by (1.3b) is nearly correct. If this requirement is satisfied, $\psi(\xi)$ can be determined by solving the equation (1.5). When T_1 is ignored the approximate Hamiltonian of the extracted system is given by

$$H_{\xi} = T(\pi, \xi) + V(\xi) + U(\xi) \quad (1.6)$$

where T is the kinetic energy and U is the potential of the average field due to the skeleton. In this case it may be natural to take into account the effect of T_1 as a perturbation. If the total system is an isolated system of particles and if the partial ξ -system contains a number of particles, $V(\xi)$ represents the interaction between particles in the extracted system. In such a case V and U are usually given in the form

$$V = \sum_j J(\xi_j - \xi_k), \quad U = \sum_k G(\xi_k). \quad (1.7)$$

The present method can equally be applied to the nuclear, atomic and molecular systems.

A practical way of determining the potential $U(\xi)$ is to make use of the variational method in which the function $\phi(\xi, x)$ is determined so as to minimize the variation integral for every value of ξ :

$$U(\xi) = \int \phi^*(\xi, x) [H_0(p, x) + H_1(0, \xi; p, x)] \phi(\xi, x) dx. \quad (1.8)$$

The adiabatic potential between atoms is often calculated in this way⁷. More simply we can calculate $U(\xi)$ by assuming a suitable function for $\phi(\xi, x)$ in this integral. For example, if the ξ -system consists of valence electrons of an atom we can replace the self-consistent function of the core for $\phi(\xi, x)$ which is then eventually independent of ξ . In the extreme case the x -system is frozen, namely the integral (1.8) is calculated by assuming that $p_1 = p_2 = p_3 = \dots = p_{N-n} = 0$ and $\phi^* \phi = \delta(x_1 - a_1) \delta(x_2 - a_2) \dots \delta(x_{N-n} - a_{N-n})$ where a_1, a_2, \dots, a_{N-n} are constants. This is the case with the nucleus in the consideration of the electronic states of an atom. A short cut is to assume a suitable, functional form of $U(\xi)$ on a certain physical ground. The free-electron model for the π -electron system is its example.

How to separate the partial Hamiltonian is most important. This procedure is not unique. Whether the method succeeds or not is due mainly to the way of this separating procedure. An ingenious device of dividing the total Hamiltonian is often capable of absorbing the greater part of the effect of $V(\xi)$ into $U(\xi)$. For this reason the particle interaction is not necessarily the same as the one contained in the original Hamiltonian. The way of the division should be devised so as to make the succeeding calculation as easy as possible or to make the effect of V as small as possible. In the most fortunate case we have an independent-particle model.

§ 2. Practical application

In this section the application of the general method mentioned in the preceding section to the practical problem will briefly be accounted for in order to obtain the better understanding of the actual meaning of the method. To begin with the simplest problem we first take up the electronic state of the helium atom in which the nucleus is frozen, and examine whether the requirement for the smallness of the effect of T_1 is satisfied.

The Hamiltonian of the helium-like atom is given by (in atomic units)

$$H = -(J_1 + J_2)/2 - Z/r_1 - Z/r_2 + 1/r_{12}. \quad (2.1)$$

In this case the Hamiltonians of partial systems are $-\Delta_1/2 - Z/r_1$ and $-\Delta_2/2 - Z/r_2$ respectively and the interaction between them is $1/r_{12}$ which contains no momentum. If $\Psi(\mathbf{x}_1, \mathbf{x}_2)$ is a solution of $(H-E)\Psi=0$ any linear combination of $\Psi(\mathbf{x}_1, \mathbf{x}_2)$ and $\Psi(\mathbf{x}_2, \mathbf{x}_1)$ also satisfies the same equation. Since $\Psi(\mathbf{x}_1, \mathbf{x}_2) + \Psi(\mathbf{x}_2, \mathbf{x}_1)$ and $\Psi(\mathbf{x}_1, \mathbf{x}_2) - \Psi(\mathbf{x}_2, \mathbf{x}_1)$ are the orbital functions for the singlet and triplet states respectively we have only to seek $\Psi(\mathbf{x}_1, \mathbf{x}_2)$. According to the procedure (1.2) we first consider the following equation:

$$(-\Delta_1/2 - Z/r_1 + 1/r_{12})\phi(\mathbf{x}_1, \mathbf{x}_2) = U\phi(\mathbf{x}_1, \mathbf{x}_2) \quad (2.2)$$

where U is an eigenvalue depending on \mathbf{x}_2 .

We want to have $\Psi(\mathbf{x}_1, \mathbf{x}_2)$ corresponding to the $1snl$ configuration and let $\phi(\mathbf{x}_1, \mathbf{x}_2)$ have the s symmetry for the space rotation. For this purpose we expand the interaction and the wave function in the Legendre series as follows:

$$\begin{aligned} \phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{k=0}^{\infty} u_k(r_1, r_2) P_k(\cos \theta_{12}) \\ 1/r_{12} &= \sum_{\lambda=0}^{\infty} a_{\lambda}(r_1, r_2) P_{\lambda}(\cos \theta_{12}), \end{aligned} \quad (2.3)$$

where θ_{12} is the angle between \mathbf{x}_1 and \mathbf{x}_2 . If we substitute these series for ϕ and $1/r_{12}$ in (2.2) we have the following set of equations for $u_k(r_1, r_2)$'s on account of the linear independence of the Legendre polynomials:

$$\begin{aligned} \left[\frac{1}{2r_1} \frac{\partial^2}{\partial r_1^2} r_1 - k(k+1)/(2r_1^2) + Z/r_1 + U(r_2) \right] u_k(r_1, r_2) \\ = \sum_{s=0}^{\infty} W(k, s; r_1, r_2) u_s(r_1, r_2) \end{aligned} \quad (2.4)$$

for $k=0, 1, 2, \dots$ where

$$W(k, s; r_1, r_2) = (2k+1) \sum_{\alpha=1}^{\beta} b(k, s, \gamma+2\alpha) a_{\gamma+2\alpha}(r_1, r_2), \quad \gamma = |k-s|, \gamma+2\beta = k+s \quad (2.5)$$

$$b(l, m, n) = (1/2) \int_{-1}^1 P_l(x) P_m(x) P_n(x) dx$$

which appear when we expand the product of ϕ and $1/r_{12}$ in the Legendre series. We see from (2.4) that the eigenvalue, U , depends only on r_2 , so that we write explicitly $U(r_2)$ in (2.4). $W(k, s; r_1, r_2)$ may be interpreted as a kind of the coupling energy between $u_k P_k$ and $u_s P_s$ states.

As is well known the expansion coefficients in (2.3) and (2.5) are given by

$$\begin{aligned} a_{\lambda}(r_1, r_2) &= r_1^{\lambda} / r_2^{\lambda+1} \quad \text{for } r_1 < r_2 \\ &= r_2^{\lambda} / r_1^{\lambda+1} \quad \text{for } r_1 > r_2 \end{aligned}$$

$$b(l, m, n) = \frac{1 \cdot 3 \cdot 5 \cdots (l+m-n-1)}{2 \cdot 4 \cdot 6 \cdots (l+m-n)} \cdot \frac{1 \cdot 3 \cdot 5 \cdots (m+n-l-1)}{2 \cdot 4 \cdot 6 \cdots (m+n-l)}.$$

$$\frac{1 \cdot 3 \cdot 5 \cdots (n+l-m-1)}{2 \cdot 4 \cdot 6 \cdots (n+l-m)} \cdot \frac{2 \cdot 4 \cdot 6 \cdots (l+m+n)}{1 \cdot 3 \cdot 5 \cdots (l+m+n+1)} \quad (2.6)$$

for $l+m \geq n$, $m+n \geq l$, $n+l \geq m$, and $l+m+n = \text{even}$

$$b(l, m, n) = 0 \quad \text{otherwise.}$$

For $k=0$ and $k=1$ the equation (2.4) reduces to

$$\left[\frac{1}{2r_1} \frac{\partial^2}{\partial r_1^2} r_1 + Z/r_1 - a_0(r_1, r_2) + U(r_2) \right] u_0(r_1, r_2) \\ = \sum_{s=1}^{\infty} a_s(r_1, r_2) u_s(r_1, r_2) / (2s+1), \quad (2.7a)$$

$$\left[\frac{1}{2r_1} \frac{\partial^2}{\partial r_1^2} r_1 - 1/r_1^2 + Z/r_1 - a_0(r_1, r_2) - (2/5)a_2(r_1, r_2) + U(r_2) \right] u_1(r_1, r_2) \\ = a_1(r_1, r_2) u_0(r_1, r_2) + \sum_{s=2}^{\infty} \frac{3u_s(r_1, r_2)}{2s+1} \left[\left(\frac{s}{2s-1} \right) a_{s-1}(r_1, r_2) + \left(\frac{s+1}{2s+3} \right) a_{s+1}(r_1, r_2) \right]. \quad (2.7b)$$

In order to determine u_0 and u_1 in an approximate way we assume that the main part of u_0 is the s orbital of the lowest energy and u_1, u_2, \dots are small corrections. Since these corrections arise from $1/r_2$ in (2.2), the summands under the \sum -signs in the r.h.s. of (2.7a) and (2.7b) are the correction terms of respective equations. In order to have the first approximation we omit these terms altogether. Then we have a homogeneous equation for u_0 and an inhomogeneous equation for u_1 . These equations still contain the correction terms, namely a_0 and a_2 terms in the l.h.s., the former being more important. We replace the former by $1/r_2$ and ignore $a_0(r_1, r_2) - 1/r_2$ and $a_2(r_1, r_2)$. Since we are seeking $\Psi(\mathbf{x}_1, \mathbf{x}_2)$ corresponding to the $1s nl$ configuration the ignorance of $a_0(r_1, r_2) - 1/r_2$ is nearly correct as it vanishes for $r_1 \sim r_2$ where the wave function is large. In this approximation we readily have the solution of (2.7) as follows:

$$U(r_2) = -Z/2 + 1/r_2, \quad u_0 = \sqrt{Z^3/\pi} N(r_2) e^{-Zr_1} \\ u_1(r_1, r_2) = \sqrt{Z/(4\pi)} N(r_2) \left[(3/2) \{e^{\rho_1} - e^{-\rho_1} (1 + 2\rho_1 + 2\rho_1^2)\} (1 + \rho_2)^2 e^{-2\rho_2} / (\rho_1 \rho_2)^2 \right. \\ \left. - (2\rho_1 + \rho_1^2) e^{-\rho_1} / \rho_2^2 \right] \\ \text{for } \rho_1 < \rho_2 \quad (2.8) \\ = \sqrt{Z/(4\pi)} N(r_2) \left[\{\rho_2^3 + (3/2) (1 - \rho_2^2 - [1 + \rho_2]^2 e^{-2\rho_2})\} (1 + 2\rho_1 + 2\rho_1^2) e^{-\rho_1} \right. \\ \left. / (\rho_1 \rho_2)^2 - 2\rho_2^{-\rho_1} \right] \quad \text{for } \rho_1 > \rho_2$$

where $N(r_2)$ is the normalization constant, $\rho_1 = Zr_1$, and $\rho_2 = Zr_2$. The function u_0 is the $1s$ orbital of the hydrogen-like atom with the nuclear charge Z . The Hamiltonian of the second partial system defined by (1.6) is given by $H_2 = -\Delta_2/2 - Z/r_2 + U(r_2) = -\Delta_2/2 - (Z-1)/r_2 - Z/2$. Therefore its eigenfunction, $\psi(\mathbf{x}_2)$, is the nl orbital of the

hydrogen-like atom with the nuclear charge $Z-1$. The first approximation of the searched solution is given by

$$\mathcal{P}(\mathbf{x}_1, \mathbf{x}_2) = \phi(\mathbf{x}_1, \mathbf{x}_2) \psi(\mathbf{x}_2) = \{u_0(r_1) + u_1(r_1, r_2) \cos \theta_{12}\} \psi(\mathbf{x}_2). \quad (2.9)$$

The more exact solution and eigenvalue-function, $U(r_2)$, can be obtained by reviving some of the neglected terms or by proceeding to the higher stages of the successive approximation. It is also possible to apply the variation method. The details of such calculations will later be reported elsewhere. The correction term including $\cos \theta_{12}$ in (2.9) means the angular deformation of the 1s orbital by the electrostatic influence of the nl orbital. This can be interpreted as the influence of the pd and other sp configurations. The more exact solution includes the radial and angular deformations of higher order or the influence of higher configurations.

Now we shall examine whether the effect of $T_1(\pi, \xi)$, which appeared in (1.5a), is small in the present case. We saw that if the contribution of $T_1(\pi, \xi)$ to the eigenvalue of the total Hamiltonian is small, the basic assumption, (1.3b), is nearly correct. In the present case $T_1(\pi, \xi)$ is given by

$$T_1(\mathbf{p}_2, \mathbf{x}_2) \psi(\mathbf{x}_2) = -2^{-1} \int \phi(\mathbf{x}_1, \mathbf{x}_2) [\mathcal{A}_2 \phi(\mathbf{x}_1, \mathbf{x}_2) - \phi(\mathbf{x}_1, \mathbf{x}_2) \mathcal{A}_2] d(\mathbf{x}_1) \psi(\mathbf{x}_2). \quad (2.10a)$$

Since $\phi(\mathcal{A}_2 \phi - \phi \mathcal{A}_2) \psi = \psi' \phi \mathcal{A}_2 \phi + (\Gamma_2 \psi') (\Gamma_2 \phi^2)$, and ϕ is real and normalized for every value of \mathbf{x}_2 as the function of \mathbf{x}_1 the integral is reduced to

$$T_1(\mathbf{p}_2, \mathbf{x}_2) \psi(\mathbf{x}_2) = -2^{-1} \psi(\mathbf{x}_2) \int \phi(\mathbf{x}_1, \mathbf{x}_2) \mathcal{A}_2 \phi(\mathbf{x}_1, \mathbf{x}_2) d(\mathbf{x}_1). \quad (2.10b)$$

If we again make use of $\phi \mathcal{A}_2 \phi = \mathcal{A}_2 \phi^2 / 2 - (\Gamma_2 \phi)^2$ and the normalization property of $\phi(\mathbf{x}_1, \mathbf{x}_2)$ we have

$$T_1(\mathbf{p}_2, \mathbf{x}_2) = (1/2) \int [\Gamma_2 \phi(\mathbf{x}_1, \mathbf{x}_2)]^2 d(\mathbf{x}_1). \quad (2.11)$$

Therefore T_1 is the positive function of \mathbf{x}_2 and independent of \mathbf{p}_2 .

The contribution of T_1 to the eigenvalue, E , of the total Hamiltonian is given by

$$E_1 = \int \psi^*(\mathbf{x}_2) T_1 \psi(\mathbf{x}_2) d(\mathbf{x}_2) \quad (2.12)$$

which has the anti-binding effect as T_1 is positive. If we make use of (1.5a) we have

$$E_1 = E - \int \psi^*(\mathbf{x}_2) [-\mathcal{A}_2 / 2 - Z/r_2 + U(r_2)] \psi(\mathbf{x}_2) d(\mathbf{x}_2). \quad (2.13)$$

The numerical values of the integral were computed by Bethe⁽⁹⁾ in an approximate way in which $U(r_2)$ was calculated more exactly than the one given by (2.8). We can estimate E_1 if we subtract his value from the observed E . The results are $|E_1/E| = 10^{-6}$ for He 3D and $|E_1/E| = 6 \times 10^{-7}$ for He 4D.

Next we briefly consider the case of a complex atom. The method illustrated in

the case of helium can be generalized to this case. The total number of electrons is denoted by N . We ignore the state of the nucleus as before. The system of atomic electrons is divided into two partial systems which consist of N_0 and N electrons respectively where $N_0 = 2, 4, 10, 12, 18, 28, \dots = \sum_{l=0}^{\infty} 2(2l+1)$. According to the procedure (1.1), the total Hamiltonian is divided into three parts:

$$\begin{aligned} H &= H_v + H_0 + H_1 & H_v &= -\sum_{k=1}^N (J_k/2 + Z/r_k) + \sum_{l=1}^N \sum_{j=k+1}^N 1/r_{kj} \\ H_1 &= \sum_{k=1}^N \sum_{a=N+1}^N 1/r_{ka} & H_0 &= -\sum_{a=N+1}^N (J_a/2 + Z/r_a) + \sum_{a=N+1}^N \sum_{b=a+1}^N 1/r_{ab}. \end{aligned} \quad (2.11)$$

If $\Psi(x_1, x_2, \dots, x_N)$ is an eigenfunction of H belonging to the eigenvalue E , the function given by $\sum_P c_P P \psi$ is also an eigenfunction of H belonging to the same eigenvalue, where P denotes the permutation of coordinates, c_P is the arbitrary coefficient which may or may not depend on the electron spins, and the summation should be extended over all possible permutations. Therefore if ψ is a solution of (1.2) including the spin functions of N_0 electrons we can write the eigenfunction of the total Hamiltonian in the form

$$\Psi = \sum_P c_P P \psi(\xi) \phi(\xi, x) \quad (2.12)$$

where ξ denotes (x_1, x_2, \dots, x_N) , x denotes $(x_{N+1}, x_{N+2}, \dots, x_N)$, and the coefficient c_P should contain the spin functions of N electrons so as to satisfy the Pauli principle. The equation (1.2) may be solved in the way that the first approximation of $\psi(\xi, x)$ is the Hartree-Fock function corresponding to the solution of the equation, $H_{v,0}\psi = U_{v,0}\psi$. In this case the first approximation of $\phi(\xi, x)$ represents a closed shell structure. Therefore one may speak that the atom consists of N valence electrons and a closed shell core of N_0 electrons, and $U(\xi)$ is the potential energy of valence electrons due to the atomic core. The deviation of exact $\phi(\xi, x)$ from the Hartree-Fock function represents the deformation of the core function due to the electrostatic influence of valence electrons or it can be interpreted as the configuration interactions. The π electron system and the skeleton of a conjugated molecule can be considered quite in the same way.

§ 3. Linear model

In this section we consider the linear model. In order to make the discussion clear-cut we take up the π -electron system, although the consideration is equally applicable to other system. The partial system is separated in two steps. First the π -electron system is extracted from a conjugated molecule where ξ -coordinates correspond to the three-dimensional cartesian coordinates of π -electrons. We next divide the π -electron system into two parts, the x and y systems. Although the interaction between π -electrons is not necessarily Coulombian as was mentioned in the previous section, we derive the one-dimensional two-electron interaction in the extracted system by assuming the Coulomb interaction between

π -electrons in order to avoid the unnecessary confusion. If the Fourier coefficient of the Coulomb potential is given by $F(k_1, k_2, k_3)$, the interaction (measured in atomic units) between electrons can be divided into three parts as follows:

$$\begin{aligned} 1/r_{12} &= J(x_1 - x_2) + G(\rho_{12}) + G_1(\mathbf{x}_1 - \mathbf{x}_2), \\ \rho_{12}^2 &= (\gamma_1 - \gamma_2)^2 + (z_1 - z_2)^2, \quad \mathbf{x}_i = (x_i, y_i, z_i), \quad i=1, 2 \end{aligned} \quad (3.1)$$

where

$$\begin{aligned} J(x) &= \sum_{k_1} F(k_1, 0, 0) \exp(i k_1 x) \\ G(\rho_{12}) &= \sum_{k_2, k_3} F(0, k_2, k_3) \exp[i[k_2(\gamma_1 - \gamma_2) + k_3(z_1 - z_2)]] \quad k_2^2 + k_3^2 > 0 \\ G_1(\mathbf{x}) &= \sum_{\mathbf{k}} F(k_1, k_2, k_3) \exp(i \mathbf{k} \cdot \mathbf{x}) \quad |\mathbf{k}| > 0, \quad k_2^2 + k_3^2 > 0 \\ \mathbf{k} &= (k_1, k_2, k_3), \quad k_j = 2\pi n_j / L_j, \quad n_j = 0, \pm 1, \pm 2, \dots, \quad j=1, 2, 3. \end{aligned} \quad (3.2)$$

The Fourier series converges to the Coulomb potential in the domain given by $2|x_1 - x_2| < L_1$, $2|\gamma_1 - \gamma_2| < L_2$, and $2|z_1 - z_2| < L_3$. We consider $J(x_1 - x_2)$ as the two-electron interaction in the partial x -system, $G(\rho_{12})$ as that in the γz -system, and $\sum_{i=1}^n \sum_{j=i+1}^n G_1(\mathbf{x}_i - \mathbf{x}_j)$ as the interaction between these two systems where n is the number of π electrons.

In order to see how $J(x)$ depends on x , we have to evaluate the sum of the series. This can easily be carried out because $J(x)$ is the average of the Coulomb potential in the γz domain. Therefore we have

$$J(x) = (4ab)^{-1} \int_{-b}^b dy \int_{-a}^a r^{-1} dz \quad (3.3)$$

where $r^2 = x^2 + y^2 + z^2$, $r \geq 0$, $2a = L_3$ and $2b = L_2$. This integral can be evaluated in an elementary way. The result is given by

$$J(x) = \frac{1}{2} \log \left(\frac{\sqrt{c^2 + x^2} + a}{\sqrt{c^2 + x^2} - a} \right)^{1/a} \left(\frac{\sqrt{c^2 + x^2} + b}{\sqrt{c^2 + x^2} - b} \right)^{1/b} - \frac{x}{ab} \tan^{-1} \frac{ab}{x \sqrt{c^2 + x^2}} \quad (3.4)$$

where $c^2 = a^2 + b^2$. The asymptotic form of this function is $1/|x|$ as is expected. If we denote the difference of $1/|x|$ and $J(x)$ by $J_0(x)$ we can divide the Coulomb potential into the sum of these two functions:

$$1/|x| = J_0(x) + J(x). \quad (3.5)$$

The value of $J(x)$ at the origin is given by $J(0) = [\log(3 + 2\sqrt{2})]/a = 1.76/a$ when a and b are the same. Therefore the most part of the Coulomb potential is reserved in $J_0(x)$ near the origin and in $J(x)$ at the distant point.

The function given by (3.4) is too complicated to see the qualitative behaviour of $J(x)$. The behaviour can easily be seen if we evaluate the sum in an approximate way by taking the average over the circular domain restricted by $y^2 + z^2 \leq a^2$, instead of the

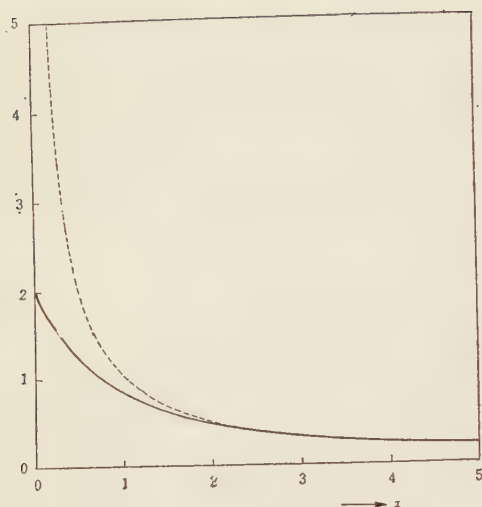


Fig. 1.
—: $J(x)$ given by (3.6)
.....: Coulomb potential

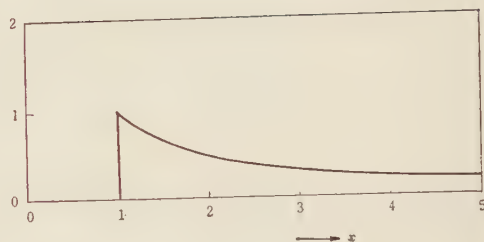


Fig. 2.
 $J(x)$ given by (3.7) ($a=1$)

rectangular domain. We have then*

$$J(x) = (2/a^2) [\sqrt{a^2 + x^2} - |x|]. \quad (3.6)$$

Its value at the origin is somewhat larger than the former case, namely $J(0) = 2/a$, because of the smaller area of the domain. It may be natural to refer $J_0(x)$ and $J(x)$ as the short range and long range parts of the Coulomb potential respectively. The division given by (3.5) depends on the value of a . Therefore a is the division constant which discriminates two parts and its genetic history illustrated above has no significance so far as the exact functional form itself of $J(x)$ at the short distance is unimportant for the final result. It is then only important that $J(x)$ is the long range part of the Coulomb potential. Another functional form is equally possible in this case. For example, the following function has the same property:

$$\begin{aligned} J(x) &= 0 \quad \text{for } |x| < a \\ &= 1/|x| \quad \text{for } |x| \geq a. \end{aligned} \quad (3.7)$$

The functional form of $J(x)$ given by (3.4) was applied to the one-dimensional free-electron model of carotenoids¹⁾ and cyanins²⁾ and the very good agreement between theory and experiment was obtained. In the former case the value of the division constant a was equal to $7 \cdot 12 \text{ \AA}$ ($a=b$), and in the latter case it was equal to $16 \cdot 7 \text{ \AA}$. The greater value of the constant means that the effect of two-electron interaction is absorbed

* If the interaction is the Yukawa potential, $e^{-\kappa r}/r$, we have

$$J(x) = \{2/(\kappa a^2)\} [\exp(-\kappa|x|) - \exp\{-\kappa\sqrt{a^2 + x^2}\}]$$

the more in the skeleton field. Nevertheless the latter value is still too large compared with the total length of the conjugated chain, while the former value is reasonable. Although, in the first paper, a^2 was interpreted as the sectional area of the pipe in which π electrons are enclosed, the interpretation may not be adequate but the present idea be better. This will further be justified by the example in the next section. However, another interpretation is still possible^(2,4).

When we extract the two-dimensional partial system we can consider $G(\rho_{12}) + F(0, 0, 0)$ as the two-electron interaction in the extracted system which is given by

$$G(\rho) + F(0, 0, 0) = \frac{1}{2a} \int_{-a}^a r^{-1} dz = \frac{1}{2a} \log \frac{\sqrt{a^2 + \rho^2} + a}{\sqrt{a^2 + \rho^2} - a}, \quad \rho^2 = x^2 + y^2. \quad (3.8)$$

This function asymptotically behaves as $1/\rho$. Still it has a weaker logarithmic singularity at $\rho=0$. This interaction was examined in the two-dimensional free-electron model of naphthalene by Huzinaga⁽⁴⁾.

§ 4. Ring model

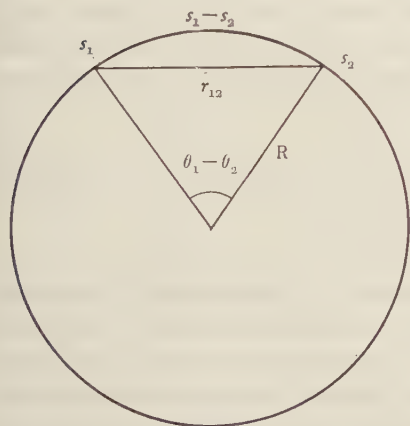


Fig. 3.

We next consider the ring model in which electrons are confined to move freely in a ring by means of suitably assumed form of $U(\xi)$. We want to derive a one-dimensional interaction in the similar way as the preceding case. Let the radius of the ring be R . The one-dimensional coordinate of a point on the ring can be given by the arc length between the point and a definite point on the ring where the coordinate and the definite point are referred to as the ring coordinate and its origin respectively. We denote the ring coordinate by s . Let the positions of two electrons be denoted by (ρ_1, θ_1, z_1) and (ρ_2, θ_2, z_2) in the cylindrical coordinates, where $\rho^2 = x^2 + y^2$ and θ is the azimuthal angle.

The distance, r_{12} , between these electrons is given by

$$r_{12}^2 = (\rho_1 - \rho_2)^2 + \rho_1 \rho_2 [2 \sin \{(\theta_1 - \theta_2)/2\}]^2 + (z_1 - z_2)^2. \quad (4.1)$$

We define a function of ρ , θ and z by

$$W(\rho, \theta, z; R^2) = [\rho^2 + (2R \sin \theta/2)^2 + z^2]^{-1/2}. \quad (4.2)$$

If $F(k_1, n, k_2; R^2)$ is the Fourier coefficient of this function we can separate its θ -part as was done in the previous section. If we replace θ by s/R , the θ -part of the function is given by

$$f(s) = \sum_{n=0}^{\infty} F(0, n, 0; R^2) e^{in s/R}. \quad (4.3)$$

By making use of the same Fourier coefficient we can expand the Coulomb interac-

tion between two electrons in the exponential series as follows:

$$\begin{aligned} 1/r_{12} &= W(\rho_1 - \rho_2, \theta_1 - \theta_2, z_1 - z_2; \rho_1, \rho_2) \\ &= \sum_{k_1, k_2, n} F(k_1, n, k_2; \rho_1, \rho_2) \exp i[k_1(\rho_1 - \rho_2) + n(\theta_1 - \theta_2) + k_2(z_1 - z_2)] \quad (4.4) \\ k_i &= 2\pi n_i / L_i, \quad i=1, 2; \quad n_1, n_2, n=0, \pm 1, \pm 2, \dots \end{aligned}$$

The function $f(s_1 - s_2)$ can be obtained by replacing R^2 for ρ_1, ρ_2 and $(s_1 - s_2)/R$ for $\theta_1 - \theta_2$ in the θ -series part of this expansion. Therefore we can consider the two-electron function given by $f(s_1 - s_2)$ as the s -part (on the ring) of the Coulomb interaction between two electrons. The sum of the Fourier series in the right-hand side of (4.3) can be evaluated by taking the average of $W(\rho, \theta, z; R^2)$ with respect to ρ and z over the domain defined by $2|\rho| \leq L_1$, and $2z \leq L_2$. The result is given by the similar expression as (3.4). As was done in the previous section, if we approximately calculate the average over the circular domain restricted by $\rho^2 + z^2 \leq a^2$, we have

$$f(s) = (2/a)^2 [\sqrt{a^2 + |2R \sin \{s/(2R)\}|^2} - |2R \sin \{s/(2R)\}|]. \quad (4.5)$$

Since the distance between two electrons on the ring is equal to $|2R \sin(s_1 - s_2)/(2R)|$ as is seen from the equation (4.1) by putting $z_1 = z_2$ and replacing R for ρ_1 and ρ_2 , the above equation has the same meaning as (3.6). Therefore the function can be replaced by

$$\begin{aligned} f(s) &= 0 && \text{for } s < s_0 \\ &= |2R \sin \{s/(2R)\}|^{-1} && \text{for } s \geq s_0. \end{aligned} \quad (4.6)$$

Making use of this form of the two-electron interaction the energy levels of π -electron systems were calculated by assuming a free-electron model for polyacenes, a preliminary account of which was already given¹¹. The length of the division constant, s_0 , was $0.68l$ where $l=1.4\text{\AA}$ is the bond length in benzene. We saw that the variation of the wave length of the absorption maxima with the number of benzene rings was reproduced fairly well. The detail of the calculation and the application of the same method to the energy levels of benzene by taking into account the configuration interaction will be reported in the separate papers. We shall see that the result is fairly good in spite of the simpleness of the model and the calculation.

§ 5. Plasma oscillation of π electrons in carotenoids

The excitation energy of the π -electron system of carotenoids was calculated by assuming the linear model in a previous paper¹. The result is given by

$$JE = \sum_{\nu=-\infty}^{\infty} N_{\nu} \sqrt{T_{\nu}^2 + 2T_{\nu} U_{\nu}} \quad (5.1)$$

where N_{ν} is the number of sound Bosons whose energy is equal to $\sqrt{T_{\nu}^2 + 2T_{\nu} U_{\nu}}$. T_{ν} and U_{ν} come from the kinetic energy and the interaction between π electrons respectively and they are given by

$$T_\nu = (h/L)^2 \cdot (n_0/m) |\nu|, \quad U_\nu = 2|\nu| J_\nu, \quad (5.2)$$

where L is the length of the conjugated chain, m is the electron mass, n_0 is the Fermi maximum, and J_ν is the Fourier coefficient of the one-dimensional interaction between π electrons. If we assume that the interaction is given by $J(x_1 - x_2)$ of (3.4) we have

$$J_\nu = \left(\frac{e^2 L}{\pi A} \right) \cdot \frac{1}{\nu^2} \quad (\nu = \pm 1, \pm 2, \dots) \quad (5.3)$$

where L_1 is replaced by L and L_2, L_3 by A . In this case the second term under the square root sign of (5.1) is independent of ν^3 .

If we denote the total number of π electrons by N we have $N = 4n_0 + 2$. When N is very large we can replace $4n_0$ with N and we have approximately

$$2T_\nu U_\nu = (\hbar\omega)^2 \quad (5.4)$$

where ω is the so-called plasma frequency⁵⁽⁶⁾:

$$\omega = \sqrt{4\pi e^2 N / (mLA)}. \quad (5.5)$$

The linear density, N/L , of π electrons is the same for all carotenoids. Therefore the plasma frequency is independent of the molecular length of carotenoids. Since T_ν is inversely proportional to the molecular length the energy of sound Bosons approaches to the limit, $\hbar\omega$, the energy of the plasma oscillation, when the molecule becomes longer and longer. The energy of low-energy (small- ν) Bosons in long carotenoids is approximately equal to the limit. The limit of the excitation energy of carotenoids is equal to

$$(\Delta E)_0 = N_R \hbar\omega \quad (5.6)$$

where N_R is the total number of sound Bosons.

As is well known, the wave length of the absorption maxima of carotenoids becomes longer and longer with the molecular length but ultimately converges to a limit. The first excitation energy of carotenoid is given by (5.1) with $N_1 = 1$ and $N_\nu = 0$ for $\nu \neq 1$. Therefore the convergence limit of the absorption energy is equal to the energy, $\hbar\omega$, of the plasma oscillation. The corresponding wave length can be calculated by employing the parameter value previously determined¹⁾, namely $A = 724.3$ atomic units. The linear density of π electrons is equal to $N/L = 1/l$ where l is the mean bond length between carbons in the conjugated chain and is assumed to be the bond length in benzene ($l = 1.4 \text{ \AA} = 2.646$ atomic units). The calculated values of the wave length and the wave number of the plasma oscillation are equal to $563 \text{ m}\mu$ and $1.78 \times 10^4 \text{ cm}^{-1}$ respectively. Whether the energy of the plasma oscillation is greater or less than the kinetic part of the absorption energy can be indicated by the ratio, $\delta = 2U_1/T_1 = (\hbar\omega/T_1)^2$. In fact the values of this ratio are large compared with unity for long carotenoids as is seen from Table 1. Therefore we see that the light absorption of long carotenoids is due mainly to the plasma oscillation of π -electrons contained in their conjugated chains. We see, from the reasoning of the third and fifth sections, that the plasma oscillation is caused by the long range part of the Coulomb interaction between

π -electrons. The screening of the short range part arises from extracting the one-dimensional partial system and making sleep the motion of π -electrons in the remaining two dimensions, while this is the charge screening which results in a cut-off of the high frequency Fourier components of the Coulomb interaction by introducing the collective coordinates in the case of the three-dimensional plasma oscillation⁵⁾⁽⁶⁾. The effective electron density of plasma oscillation may be defined as $N/(LA)=1/(lA)$ though A does not mean the real sectional dimension of the electron vessel in the present case. The calculated value of the effective density for carotenoids is equal to $3.5 \times 10^{21} \text{ cm}^{-3}$. The calculated and observed values of absorption maxima are also shown in Table 1.

Table 1. Absorption Maxima of Long Carotenoids

N : number of π electrons

λ_{calc} , λ_{obs} : wave lengths of calculated and observed absorption maxima measured in $m\mu$

λ_{calc} : wave length calculated by (5.1) with $N_1=1$ and $N_v=0$ for $v \neq 1$

$\delta = 2U_1/T_1$

| N | δ | λ_{calc} | λ_{obs}^* | Carotenoids |
|----------|----------|------------------|-------------------|---------------------------------|
| 20 | 1.47 | 457 | 450 | Anthraxanthin |
| | | | 461 | Rubixanthinmonoeoxyde |
| 22 | 1.76 | 471 | 466 | Astaxanthin |
| | | | 471 | Rhodopin |
| | | | 472 | Lycoxanthin |
| | | | 473 | Lycophyll |
| | | | 473 | Lycopene |
| 24 | 2.07 | 483 | 478 | Eschscholtzanthin |
| | | | 482 | Rhodoxanthin |
| 26 | 2.42 | 493 | 493 | Rhodoviolasin |
| | | | 495 | Decapreno- ϵ -carotene |
| 28 | 2.79 | 501 | 500 | Anhydroeschscholtzanthin |
| 30 | 3.18 | 508 | 504 | Dehydrolycopene |
| | | | 508 | Decapreno- β -carotene |
| 38 | 5.03 | 528 | 537 | Dodecapreno- β -carotene |
| ∞ | ∞ | 553 | | (plasma oscillation) |

*Sources were given in reference (1).

References

- 1) G. Araki and T. Murai, *Prog. Theor. Phys.* **8**(1952), 639;
J. Chem. Phys. **20**(1952), 1661.
- 2) G. Araki and H. Araki, *Prog. Theor. Phys.* **11** (1954), 20.
- 3) G. Araki and T. Murai, *J. Chem. Phys.* **22** (1954), 945.
- 4) S. Huzinaga, *Prog. Theor. Phys.* **15** (1956), 495.
- 5) D. Bohm and D. Pines, *Phys. Rev.* **92**(1953), 609.
- 6) S. Tomonaga, *Prog. Theor. Phys.* **13**(1954) 467, 482.
- 7) G. Araki, S. Tutihasi, and W. Watari, *Prog. Theor. Phys.* **6**(1951), 135, 945, 961.
- 8) S. Tomonaga, *Prog. Theor. Phys.* **5**(1950), 544.
- 9) H. Bethe, *Geiger-Scheel's Handb. der Phys.*, **24/1** (1933), 339.

Note added in proof. Recently Ta-You-Wu and A. B. Bhatia [*J. Chem. Phys.*, **24** (1956), 48, 444] calculated the operator T_1 in the cases of ionic and neutral hydrogen molecules and two normal helium atoms which were divided into electronic and nuclear systems. They found that T_1 is of the same order of magnitude as the van der Waals interaction and that T_1 is represented by a repulsive potential. The latter result can be seen from (2.11) in the present paper.

The estimation of $|E_1/E|$ in the present paper may include a certain error because Bethe's calculation is not sufficiently accurate for the purpose of calculating E_1 according to Equation (2.13). It is better to calculate E_1 directly by making use of Equations (2.11) and (2.12). The eigenvalue of $-A_2/2 - Z/r_2 + U(r_2)$ is lower than the eigenvalue of the Hamiltonian given by (2.1) because T_1 is repulsive as is seen from (2.11). If the former eigenvalue is calculated in an approximate way the difference of these eigenvalues is partly compensated by the error of the calculation. We consequently underestimate the value of E_1 if we make use of (2.13). Thus the estimated value of $|E_1/E|$ in the present paper may be too small. The result of direct calculation of E_1 will be reported elsewhere.

Free-Electron Model of Benzene

Gentaro ARAKI

Faculty of Engineering, Kyoto University, Yosida, Kyoto

(Received April 27, 1956)

Lower excited electronic levels of benzene are calculated by assuming π electrons to move freely in a ring whose perimeter is equal to that of the benzene molecule. The interaction between π electrons is assumed to be the long range part separated from the Coulomb interaction by the zero cut-off at $0.46l$ of the ring coordinate where l is the distance between adjacent carbons in the benzene molecule. The configuration interaction with the adjacent levels is taken into account. All interaction integrals can easily be evaluated without any neglect. The calculated result agrees fairly well with experiment in spite of simpleness of the model.

Introduction

The free-electron model succeeded in explaining the variation of absorption maxima corresponding to the change in the chain length in the case of molecules with conjugated linear chains¹⁾²⁾. If this success indicates that an essential character of the π -electron system can be represented by a free-electron model, the same must hold in the case of molecules with conjugated rings. It must be worth while to examine whether this is true because, if so, we have a further justification for understanding the essential character of π electrons as that of free electrons.

The benzene molecule is the simplest of conjugated ring molecules. Its electronic states have theoretically been studied by many authors³ according to the orthodox method, but the determination of energy levels includes a very tedious task of evaluating difficult integrals. On the contrary the free-electron model does not include such difficult integrals. Orbitals are orthogonal in this model. In the present paper lower excited electronic levels of benzene will be calculated as a first step to examine the free-electron model for ring molecules. The lowest level belonging to each symmetry will be calculated taking into account the configuration interaction with the adjacent levels of the same symmetry. The interaction between π electrons in the present ring model will be assumed to be the long range part, of the Coulomb interaction, which is obtained by a zero cut-off of its short range part. The significance of such an interaction was accounted for in the previous paper¹⁾ as well as the general method of approximation included in the present model. The application of the model to other molecules will be examined in a separate paper.

We shall see that the calculated result for singlets is in fairly good agreement with experiment. This may be considered as adding a further evidence to the free-electron

property of π electrons in the conjugated system. We may further consider this as a justification to the assumption on the interaction between the π electrons. It is not certain, however, whether the present model is wholly in agreement with experiment because

observed levels have not yet fully known on the one hand and the present model can not discriminate B_{1u} and B_{2u} on the other hand. Be that as it may, there is no evidence that the present model certainly disagrees with experiment. If we compare the present result with the LCAO MO calculation, we see that its agreement with experiment is rather better than the latter.

§ 1. Classification of eigenstates

We assume the ring model for π electrons of a benzene molecule, in which π electrons are

enclosed in a ring whose perimeter is equal to that of the benzene molecule. The position of a point on the ring is specified by an arc length from a reference point as is shown in Fig. 1. The arc length is denoted by s which is referred to as the ring coordinate of the point. If the total number of π electrons is N , the Hamiltonian of the π -electron system is given by

$$H = -\frac{1}{2} \sum_{k=1}^N \frac{d^2}{ds_k^2} + \sum_{k=1}^N \sum_{j=k+1}^N J(s_k - s_j), \quad (1.1)$$

where all quantities are measured in atomic units. The function $J(s_k - s_j)$ represents the interaction between two electrons and is defined by

$$\begin{aligned} J(s_k - s_j) &= 0 & |s_k - s_j| < s_0 \\ &= 1/r_{kj} & |s_k - s_j| \geq s_0, \end{aligned} \quad (1.2)$$

where r_{kj} is the chord length between two points specified by s_k and s_j (see Fig. 1). This interaction is the long range part of the Coulomb interaction and s_0 is the division constant which discriminates the short and long range parts.⁴⁾

We adopt the eigenfunctions of the one-electron kinetic energy as the molecular orbitals and expand the eigenfunction of H in a series of determinantal functions which consist of these orbitals and appropriate spin functions. The normalized orbitals are given by

$$\varphi_q(s) = L^{-1/2} \exp(2\pi i q s / L), \quad q = 0, \pm 1, \pm 2, \dots, \quad (1.3)$$

where L is the perimeter length of the ring and q is referred to as the one-electron ring quantum number.* If we denote the operator which displaces s to $s + L/m$ by C_m the

* Platt⁵⁾ perhaps omitted negative values of the ring quantum number whereas all integer values are included in the present paper.

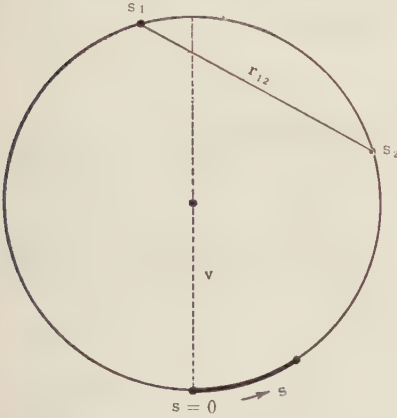


Fig. 1

orbital is the eigenfunction of this operator:

$$C_m \varphi_q = \omega^q \varphi_q \quad \omega = e^{2\pi i/m} \quad (1.4)$$

where m is the integer.

We define the ring quantum number of the total system by

$$Q = \sum_{i=1}^N q_i \quad (1.5)$$

where q_1, q_2, \dots , and q_N are the one-electron ring quantum numbers of orbitals included in the anti-symmetrical wave function of the total system. If ψ_Q is the anti-symmetrical wave function of the total system with the ring quantum number, Q , this is the eigenfunction of C_m because of (1.4). If σ_v is the operator which reflects the space with respect to the vertical plane, v , through the centre and the reference point of the ring (see Fig. 1), the operator changes the sign of Q :

$$C_m \psi_Q = \omega^Q \psi_Q; \quad \sigma_v \psi_Q = \psi_{-Q} \quad (Q \neq 0), \quad (1.6)$$

where the latter is the well-known property of the eigenfunction of C_m . Further we denote the inversion operator with respect to the centre of the ring by I . The parity of the orbital for this inversion is g or u according as its ring quantum number is even or odd as is seen from Fig. 2. Therefore the same is true for the wave function of the total system:

$$I \psi_Q = (-1)^Q \psi_Q. \quad (1.7)$$

In case of the benzene molecule the symmetry operators are C_6 , σ , and I . In this case we have $m=6$, $\omega + \omega^{-1} = 2 \cos(2\pi/m) = 1$, $\omega^2 + \omega^{-2} = 2 \cos(4\pi/m) = -1$, $\omega^3 = -1$, $\omega^4 = \omega^{-2}$, $\omega^5 = \omega^{-1}$ and $\omega^6 = 1$. We have thus the following correspondence between the ring quantum number and the state of the total system:

| | | Q | | | | | | | | | | |
|--------|---|--------|------------|-------|----------|------------|----------|-------|------------|-------|--|--|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | | (1·8) | | |
| | | States | $A_{1,2g}$ | E_u | E_{2g} | $B_{1,2u}$ | E_{2g} | E_u | $A_{1,2g}$ | | | |
| cos | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| sin | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| (q) | 0 | 1 | 2 | 3 | | | | 3 | | | | |
| parity | g | u | g | u | | | | u | | | | |

Fig. 2. The parity of orbitals

The real and imaginary parts of the orbital are indicated by cos and sin respectively.

The ring quantum number is a true quantum number for the ring model because its Hamiltonian has a circular symmetry, and the states of positive and negative ring

quantum numbers do not mix, though this is not the case with the real benzene molecule. Therefore we have no need of considering the states of negative ring quantum numbers, but unfortunately we can not discriminate the B_1 level from B_2 for the same reason. Rough values of levels are given by the eigenvalues of the kinetic energy. The one-electron kinetic energy corresponding to the molecular orbital with q is given by

$$\varepsilon(q) = 2(\pi/L)^2 q^2. \quad (1.9)$$

Therefore the eigenvalue of the kinetic energy is proportional to

$$W = \sum_{k=1}^N q_k^2. \quad (1.10)$$

Since we want to calculate the lowest level for each symmetry, we have only to consider the configurations of lower energy. Such configurations are shown in Table 1 where the numbers in the parentheses are the values of W .

Table 1

| | | | | | | | |
|-------|------------|------|--|------|---|------|---|
| $Q=0$ | $A_{1,2g}$ | (4) | $\varphi_0^2 \varphi_1^2 \varphi_{-1}^2$ | (10) | $\varphi_0^2 \varphi_1 \varphi_{-1} \varphi_2 \varphi_{-2}$ | (12) | $\varphi_1^2 \varphi_{-1}^2 \varphi_2 \varphi_{-2}$ |
| $Q=1$ | E_{1u} | (7) | $\varphi_0^2 \varphi_{-1}^2 \varphi_1 \varphi_2$ | (11) | $\varphi_1^2 \varphi_0 \varphi_{-1} \varphi_2 \varphi_{-2}$ | (13) | $\varphi_0^2 \varphi_2^2 \varphi_{-1} \varphi_{-2}$ |
| $Q=2$ | E_{2g} | (8) | $\varphi_1^2 \varphi_{-1}^2 \varphi_0 \varphi_2$ | (10) | $\varphi_0^2 \varphi_1^2 \varphi_2 \varphi_{-2}$ | (10) | $\varphi_0^2 \varphi_{-1}^2 \varphi_2^2$ |
| | | (12) | $\varphi_0^2 \varphi_{-1}^2 \varphi_1 \varphi_3$ | | | | |
| $Q=3$ | $B_{1,2u}$ | (7) | $\varphi_0^2 \varphi_1^2 \varphi_{-1} \varphi_2$ | (11) | $\varphi_{-1}^2 \varphi_2^2 \varphi_0 \varphi_1$ | (13) | $\varphi_0^2 \varphi_2^2 \varphi_1 \varphi_{-2}$ |
| | | (13) | $\varphi_1^2 \varphi_{-1}^2 \varphi_0 \varphi_3$ | | | | |
| $Q=4$ | E_{2g} | (10) | $\varphi_0^2 \varphi_2^2 \varphi_1 \varphi_{-1}$ | (12) | $\varphi_1^2 \varphi_{-1}^2 \varphi_2^2$ | (12) | $\varphi_0^2 \varphi_1^2 \varphi_{-1} \varphi_3$ |
| $Q=5$ | E_{1u} | (11) | $\varphi_1^2 \varphi_2^2 \varphi_0 \varphi_{-1}$ | (15) | $\varphi_0^2 \varphi_1 \varphi_{-1} \varphi_2 \varphi_3$ | | |
| $Q=6$ | $A_{1,2g}$ | (10) | $\varphi_0^2 \varphi_1^2 \varphi_2^2$ | (16) | $\varphi_1^2 \varphi_0 \varphi_{-1} \varphi_2 \varphi_3$ | | |

§ 2. Energy levels

Since the orbitals in the present model are orthonormal, the standard method in the theory of atomic spectra can be applied to calculating matrix elements of the Hamiltonian given by (1.1). The wave functions are the eigenfunctions of the kinetic energy of the system. Therefore we have only to calculate matrix elements of the interaction between electrons. The direct integrals are independent of orbitals because the electron density is uniform throughout the ring, and their contribution has no influence on the excitation energy. All interaction integrals arising from permutations of electrons can be expressed in terms of the following integrals as will be shown in the last section:

$$D(q) = L^{-2} \int_0^L \int_0^L J(s_1 - s_2) \exp \{2\pi i q (s_1 - s_2)/L\} ds_1 ds_2. \quad (2.1)$$

The calculation of this integral is quite easy as will be accounted for in the last section. The orthogonality of the orbitals and the easiness of calculating the interaction integrals make the theoretical aspect of the present model very simple. We have no need of evaluating the many-centre integrals or of computing extensive numerical values of integrals by the tedious task.

For the purpose of calculating the lowest energy level for the definite symmetry we first examine the diagonal elements of the Hamiltonian for various states of the given symmetry. In Table 1 the configurations are arranged in the order of the values of the diagonal elements. The first member in each row is the lowest of those with the symmetry shown on the left. The state of the lowest diagonal energy is referred to as the a-state and that of the next higher diagonal energy as the b-state. We take into account the interaction between these two states only and ignore the influence of all higher states. If the matrix elements of the Hamiltonian are denoted as

$$\alpha = (\psi_a, H \psi_a), \quad \beta = (\psi_b, H \psi_b), \quad \gamma = (\psi_a, H \psi_b) \quad (2.2)$$

the lowest energy level, E , of the given symmetry is given by

$$E = \alpha - \Delta E, \quad \Delta E = \sqrt{(\beta - \alpha)^2/4 + |\gamma|^2} - (\beta - \alpha)/2 \quad (2.3)$$

where ψ_a and ψ_b are the normalized wave functions of the lowest and next higher states respectively and ΔE represents the depression of the lowest energy level caused by its interaction with the next higher level. If more than three orbitals in ψ_b are different from those in ψ_a , the former has no influence on the latter because γ vanishes in this case. In such a case we should further take into account the influence of the next higher state. Fortunately such a case does not occur in the present consideration. Further we see from the numerical computation that the influence of the ignored states is small and the ignorance gives no serious difficulty in the present case.

When all orbitals in a configuration are closed with respect to the spin, the configuration contains only a singlet as is the case with $\varphi_0^2 \varphi_1^2 \varphi_{-1}^2 (Q=0)$. If two orbitals in a configuration are open and the others are closed the configuration still contains only one singlet and one triplet. The examples of this case are $\varphi_0^2 \varphi_{-1}^2 \varphi_1 \varphi_2 (Q=1)$, $\varphi_1^2 \varphi_{-1}^2 \varphi_0 \varphi_2 (Q=2)$ and $\varphi_0^2 \varphi_1^2 \varphi_{-1} \varphi_2 (Q=3)$. If a configuration consists of four open and one closed orbitals there are two singlets, three triplets and one quintet in this configuration.

As an example of the last case we consider the configuration $\varphi_1^2 \varphi_0 \varphi_{-1} \varphi_2 \varphi_{-2} (Q=1)$. For the sake of simplicity we represent the antisymmetrical normalized wave function in the abbreviated way as follows. For example, the quintet wave function is written as

$$(\varphi_1^2 \varphi_0^+ \varphi_{-1}^+ \varphi_2^+ \varphi_{-2}^+) = 1/\sqrt{6!} \cdot \sum_P \varepsilon_P P \varphi_1^+ (1) \varphi_1^+ (2) \varphi_0^+ (3) \varphi_{-1}^+ (4) \varphi_2^+ (5) \varphi_{-2}^+ (6) \quad (2.4)$$

where P denotes the permutation operator for electron coordinates, ε_P is equal to $+1$ or -1 according as the permutation is even or odd, and the summation is to be extended over all possible permutations. The superscript, $+$ or $-$, attached to the orbitals denotes the spin state. There are six wave functions whose spin magnetic quantum numbers, M_s , vanish and four wave functions whose spin magnetic quantum numbers are equal to unity.

They are given by

$$\begin{array}{ll}
 M_s=0 & M_s=1 \\
 \psi_1=(\varphi_1^2 \varphi_0^+ \varphi_{-1}^+ \varphi_2^- \varphi_{-2}^-) & \psi^1=(\varphi_1^2 \varphi_0^+ \varphi_{-1}^+ \varphi_2^- \varphi_{-2}^-) \\
 \psi_2=(\quad - \quad - \quad + \quad +) & \psi^2=(\quad - \quad + \quad + \quad +) \\
 \psi_3=(\quad + \quad - \quad + \quad -) & \psi^3=(\quad + \quad + \quad + \quad -) \\
 \psi_4=(\quad - \quad + \quad - \quad +) & \psi^4=(\quad + \quad + \quad + \quad +) \\
 \psi_5=(\quad + \quad - \quad - \quad +) & \\
 \psi_6=(\quad - \quad + \quad + \quad -) &
 \end{array} \quad (2.5)$$

where orbitals are not written except the first member in each set, but only one-electron spin states.

Four of the six linearly independent linear combinations of the functions in the first set belong to the quintet and the triplets. One of the four linearly independent linear combinations of the functions in the second set belongs to the quintet. We have thus two singlets and three triplets as follows:

$$\begin{array}{ll}
 S=M_s=0 & Q=1 \quad {}^1E_{1u} \\
 {}^1\psi_b=(\psi_1+\psi_2-\psi_3-\psi_4)/2 & {}^3\psi_b=(\psi^1-\psi^2)/\sqrt{2} \\
 {}^1\psi_c=(\psi_1+\psi_2-\psi_5-\psi_6)/2 & {}^3\psi_c=(\psi^3-\psi^2)/\sqrt{2} \\
 & {}^3\psi_d=(\psi^1-\psi^4)/\sqrt{2}
 \end{array} \quad (2.6)$$

where S is the total spin quantum number. From the result of the numerical calculation we see that the values of the diagonal energy are in alphabetical order of the suffices attached to ψ and the first members are lowest in each set. Therefore we designate them as the b-state. Other linear combinations are possible. For example ${}^3\psi_b-{}^3\psi_c$, ${}^3\psi_b-{}^3\psi_d$, and ${}^3\psi_b-{}^3\psi_c-{}^3\psi_d$ are all triplets, but they have higher energy. In order to avoid an unnecessary numerical work we give up to examine other possible linear combinations by solving an eigenvalue problem in each set. Such a work has little significance in the method of the present approximation. We select the b-state from $\varphi_0^2 \varphi_1 \varphi_{-1} \varphi_2 \varphi_{-2}$ in the same way.

The lowest state of the single configuration listed in Table 1 is ${}^1A_{1g}(Q=0, W=4)$. Its wave function is given by $\psi_0=(\varphi_0^2 \varphi_1^2 \varphi_{-1}^2)$. The diagonal element of H for this state is given by

$$(\psi_0, H \psi_0) = 2(\pi/L)^2 \cdot 4 + C - 4D(1) - 2D(2) \quad (2.7)$$

where C is the contribution from direct integrals. We measure the values of the diagonal elements for all states from this value:

$$E_a = (\psi_a, H \psi_a) - (\psi_0, H \psi_0), \quad E_b = (\psi_b, H \psi_b) - (\psi_0, H \psi_0). \quad (2.8)$$

Then we have $\beta - \alpha = E_b - E_a$. The wave functions and the matrix elements of the total Hamiltonian, H , given by (1.1) for various states are shown in Table 2.

Table 2.

- (i) $Q=0 \quad {}^1A_{1g}$
- $$\begin{aligned} \phi_1 &= (\varphi_0^2 \varphi_1^+ \varphi_{-1}^+ \varphi_2^- \varphi_{-2}^-) \\ \phi_2 &= (\quad - \quad - \quad + \quad) \\ \phi_3 &= (\quad + \quad - \quad + \quad -) \\ \phi_4 &= (\quad - \quad + \quad - \quad +) \\ \phi_b &= (\phi_1 + \phi_2 - \phi_3 - \phi_4)/2 \\ E_b &= 2(\pi/L)^2 \cdot 6 + [2D(1) - D(2) + 4D(3) - D(4)]/2 \\ (\phi_a, H \phi_b) &= 2D(3) - D(1) \end{aligned}$$
- (ii) $Q=1 \quad {}^1E_{1u}$
- $$\begin{aligned} \phi_a' &= (\phi_a' - \phi_a'')/\sqrt{2} \\ \phi_b' &= {}^1\phi_b \text{ given by (2.6)} \\ \phi_a'' &= (\quad - \quad + \quad) \\ E_a &= 2(\pi/L)^2 \cdot 3 + 2D(1) - D(3) \\ E_b &= 2(\pi/L)^2 \cdot 7 + [2D(1) + 3D(2) - D(4)]/2 \\ (\phi_a, H \phi_b) &= [D(1) - 2D(2)]/\sqrt{2} \end{aligned}$$
- (ii) $Q=1 \quad {}^3E_{1u}$
- $$\begin{aligned} \phi_a &= (\varphi_0^2 \varphi_{-1}^2 \varphi_1^+ \varphi_2^+) \\ \phi_b &= {}^3\phi_b \text{ given by (2.6)} \\ E_a &= 2(\pi/L)^2 \cdot 3 - D(3) \\ E_b &= 2(\pi/L)^2 \cdot 7 + [D(1) + 3D(2) - 3D(3) - D(4)]/2 \\ (\phi_a, H \phi_b) &= -D(2)/\sqrt{2} \end{aligned}$$
- (iv) $Q=2 \quad {}^{1,3}E_{2g}$
- $$\begin{aligned} \phi_a &= (\phi_a' \mp \phi_a'')/\sqrt{2} \\ \phi_b &= (\phi_b' \mp \phi_b'')/\sqrt{2} \\ \phi_a' &= (\varphi_1^2 \varphi_{-1}^2 \varphi_0^+ \varphi_2^-) \\ \phi_a'' &= (\quad - \quad + \quad) \\ \phi_b' &= (\varphi_0^2 \varphi_1^2 \varphi_{-2}^+ \varphi_2^-) \\ \phi_b'' &= (\quad - \quad + \quad) \\ E_a &= 2(\pi/L)^2 \cdot 4 + D(1) \pm D(2) - D(3) \\ E_b &= 2(\pi/L)^2 \cdot 6 + D(1) - D(3) \pm D(4) \\ (\phi_a, H \phi_b) &= -D(1) \end{aligned}$$
- upper sign : signlet ; lower sign : triplet.
- (iv) $Q=3 \quad {}^{1,3}B_u$
- $$\begin{aligned} \phi_a &= (\phi_a' \mp \phi_a'')/\sqrt{2} \\ \phi_b &= (\phi_b' \mp \phi_b'')/\sqrt{2} \\ E_a &= 2(\pi/L)^2 \cdot 3 \pm D(3) \\ E_b &= 2(\pi/L)^2 \cdot 7 + (2 \pm 1)D(1) - 2D(3) \\ (\phi_a, H \phi_b) &= D(1) \pm D(2) \end{aligned}$$
- upper sign : singlet ; lower sign : triplet.

We next compute the values of the matrix elements by making use of the numerical values of interaction integrals which are given in the last section, and then calculate the energy levels by inserting these values in the equation (2.3). The result is shown in Table 3 where E_{kin} denotes the kinetic energy. In this calculation it is assumed that $s_0=0.64l$ and $l=1.4\text{\AA}=2.64576$ atomic units where s_0 is the division constant for the interaction potential [see (1.2)] and l is the distance between two adjacent carbons in the benzene molecule⁷⁾. The energy levels, E_{calc} , are measured from the ground level, $^1A_{1g}$. The observed values of the levels are also shown in the last column of the table⁸⁾. We see that the calculated values for singlets are in agreement with experiment. Though the calculated 3B_u is not as low as 3.8eV , this does not mean the disagreement with experiment, because $^3B_{1u}$ and $^3B_{2u}$ can not be distinguished in the present model on the one hand and Niira³⁾ showed that the calculated value of $^3B_{2u}$ is equal to 9.6eV on the other hand. If his value indicates the correct position of the level, the present result may rather be considered as in agreement with experiment.

In the $^1A_{1g}$ and $^1E_{1u}$ states the interaction between the lowest two states are large compared with those in other states. The depressions, JE , of the lowest levels in these two states are nearly equal. Therefore the $^1E_{1u}$ level is nearly equal to the lowest diagonal element, E_{ca} , for this state as is seen from Table 3. In other states the levels are higher than the diagonal elements because the depressions are smaller than that of the ground states.

Table 3

The energy is measured in eV where the energy equivalence of 1 atomic unit= 27.2100 eV is assumed.⁹⁾

| state | Q | E_{kin} | E_a | E_b | ΔE | E_{calc} | $E_{obs}^{8)}$ |
|------------|---|-----------|--------|--------|------------|------------|--|
| $^1A_{1g}$ | 0 | 0 | 0 | 10.670 | 0.576 | 0 | 0 |
| $^1E_{1u}$ | 1 | 6.394 | 7.092 | 12.515 | 0.659 | 7.009 | 7.0 |
| $^3E_{1u}$ | 1 | 6.394 | 7.858 | 14.902 | 0.179 | 8.256 | — |
| $^1E_{2g}$ | 2 | 8.525 | 7.999 | 13.091 | 0.0286 | 8.547 | — |
| $^3E_{2g}$ | 2 | 8.525 | 11.214 | 14.648 | 0.0422 | 11.748 | — |
| 1B_u | 3 | 6.394 | 4.930 | 16.699 | 0.328 | 5.179 | $\begin{cases} ^1B_{1u} & 6.2 \\ ^1B_{2u} & 4.9 \end{cases}$ |
| 3B_u | 3 | 6.394 | 7.858 | 17.465 | 0.154 | 8.281 | $\begin{cases} ^3B_{1u} & 3.8 \\ ^3B_{2u} & (9.6)^* \end{cases}$ |

* the calculated value by Niira³⁾

§ 3. Interaction integrals

In order to calculate the matrix elements of Hamiltonian we have to evaluate the interaction integrals of the following type :

$$(q, n | J | p, m) = \int_0^l \int_0^l \varphi_q^*(s_1) \varphi_n^*(s_2) J(s_1 - s_2) \varphi_p(s_1) \varphi_m(s_2) ds_1 ds_2 \quad (3.1)$$

where $J(s)$ is defined by (1.2). Since the states of different total ring quantum numbers do not interact, we have $q+n=p+m$, and we see, from (1.3), that

$$\varphi_q^*(s_1)\varphi_n^*(s_2)\varphi_p(s_1)\varphi_m(s_2) = (\sqrt{L/L^2})\varphi_{p-q}(s_1-s_2).$$

Therefore we have

$$(q, n|J|p, m) = D(p-q), \quad (p-q=n-m) \quad (3.2)$$

where $D(q)$ is given by (2.1). On account of the even property of $J(s)$ the integral can be reduced to

$$D(q) = (2/L) \int_0^{L/2} J(s) \cos(2\pi qs/L) ds. \quad (3.3)$$

Therefore $D(q)$ is the even function of q . The chord length, r , corresponding to the arc length, s , is given by

$$r = (L/\pi) \sin(\pi s/L). \quad (3.4)$$

If we substitute this in $J(s)$ defined by (1.2) and then the result in (3.3), we have

$$D(q) = (2/L) \int_0^{\pi/2} [\cos(2q\theta)/\sin\theta] d\theta, \quad \phi = \pi s_0/L. \quad (3.5)$$

If we further apply the addition theorem for the trigonometric function to the integrand of $D(q) - D(q+1)$ we can remove $\sin\theta$ from the denominator of the integrand and we have the following recurrence formula:

$$D(q) - D(q+1) = (2/L) [2/(2q+1)] \cos[(2q+1)\pi s_0/L] \quad (3.6)$$

$$D(0) = (2/L) \log \cot(\pi s_0/2L)$$

whence we have

$$D(q) = \frac{2}{L} \left[\log \cot \frac{\pi s_0}{2L} - \sum_{m=0}^{q-1} \frac{2}{2m+1} \cos\left\{(2m+1) \frac{\pi s_0}{L}\right\} \right]. \quad (3.7)$$

Table 4

| s_0/l | 0.62 | 0.64 | 0.66 | 0.68 |
|--------------|---------|---------|---------|---------|
| $D(1)$ | -0.2954 | -0.3829 | -0.4665 | -0.5463 |
| $D(2)$ | -1.5800 | -1.6075 | -1.6299 | -1.6473 |
| $D(3)$ | -1.5082 | -1.4642 | -1.4154 | -1.3622 |
| $D(4)$ | -0.8760 | -0.7789 | -0.6807 | -0.5820 |
| ${}^1A_{1g}$ | 0 | 0 | 0 | 0 |
| ${}^1E_{1u}$ | 7.286 | 7.009 | 6.740 | 6.478 |
| ${}^3E_{1u}$ | 8.385 | 8.256 | 8.124 | 7.990 |
| ${}^1E_{2g}$ | 8.792 | 8.547 | 8.305 | 8.068 |
| ${}^3E_{2g}$ | 11.946 | 11.748 | 11.541 | 11.327 |
| 1B_u | 5.255 | 5.179 | 5.106 | 5.040 |
| 3B_u | 8.388 | 8.281 | 8.172 | 8.057 |

In the present case of benzene L is equal to $6l$ where l is the distance between adjacent carbons. The observed value of l is equal to 1.4\AA ($=2.64576$ atomic units)⁷⁾. We compute the numerical values of $D(q)$ for four values of s_0 . The result is shown in Table 4 where units are converted according to DuMond and Cohen's determination⁹⁾ that 1 atomic units of energy is equal to 27.2100eV. The energy values in Table 3 are computed for $s_0=0.64l$. For other values of s_0 we have different values of energy levels. They are shown in the same table in order to see this variation in energy. As is seen in the table the triplet is higher than the singlet of the same symmetry. The reason for this is that the exchange integral, $D(q)$, is negative in the present model in contrast with the atomic case because the interaction is not Coulombian.

References

- 1) G. Araki and T. Murai, Prog. Theor. Phys. 8(1952), 639.
- 2) G. Araki and H. Araki, Prog. Theor. Phys. 11(1954), 20.
- 3) M. Goeppert-Mayer and A. L. Sklar, J. Chem. Phys. 6(1938), 645.
C. C. J. Roothaan and R. G. Parr, J. Chem. Phys. 17(1949), 1001.
R. G. Parr, D. P. Craig and I. G. Ross, J. Chem. Phys. 18 (1950), 1561
K. Niira, J. Phys. Soc. Japan, 8(1953), 630.
- 4) G. Araki, Prog. Theor. Phys. in this issue.
- 5) J. R. Platt, J. Chem. Phys. 17(1949), 484.
- 6) E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra*, p. 169 (1935)
- 7) Landolt-Börnstein, *Zahlenwert und Funktionen* Bd. 1, Tl. 2(1951).
- 8) H. Sponer, G. Nordheim, A. L. Sklar and E. Teller, J. Chem. Phys. 7(1939), 207
H. Sponer, J. Chem. Phys. 8(1940), 705.
C. C. J. Roothaan and R. S. Mulliken, J. Chem. Phys. 16 (1948), 118.
- 9) J. W. M. DuMond and E. Richard Cohen, Rev. Mod. Phys. 25(1953), 691.

Note added in proof. Recently Ham and Ruedenberg attempted to introduce the electronic interaction into the free-electron network model. [N. S. Ham and K. Ruedenberg, J. Chem. Phys. 25 (1956), 1] They assigned the excitation energy of 4.4 eV and 4.9 eV to ${}^3E_{1u}$ and ${}^3B_{2u}$ respectively. If the assignment is correct for ${}^3E_{1u}$ the present result is certainly incorrect. The assignment to ${}^3B_{2u}$ is not necessarily in contradiction to the present result because the theory can reproduce the experiment if the equations (v) in Table 2 are modified by the additional terms so as to discriminate B_{1u} and B_{2u} . In order that the singlet is higher than the triplet the value of $D(q)$ must be positive. If we determine the values of $D(1)$ and $D(3)$ so that singlet-triplet energy intervals agree with the assignment by Ham and Ruedenberg and compute the excitation energy for ${}^1E_{1u}$ and 1B_u by making use of this determination the result is too large. Further if we make $D(q)$ positive by the zero cutoff at a shorter distance the situation is not improved. Ham and Ruedenberg modified the way of cutoff. They employed the straight cutoff in order to obtain the positive $D(q)$. The situation is still qualitatively the same.

The Energy Band Structure of the Metallic Copper

—The Orthogonalized Plane Wave Method—

Mitsuru FUKUCHI

Institute of Science and Technology, University of Tokyo, Tokyo

(Received May 24, 1956)

The electronic band structure in the metallic copper is worked by the orthogonalized plane wave method in which the crystal potential is constructed by using both the Hartree's self-consistent field of Cu^+ ion cores with exchange and conduction electrons with the Slater's approximate exchange potential.

The approximate energy versus \vec{k} curve is obtained, which shows the strong resemblance to the free electron behaviour along three directions in \vec{k} space, i.e., (1.0.0), (1.1.1) and (1.1.0). The results of the various methods of approximations performed so far are compared with ours together with some critical discussions.

§ 1. Introduction

As is well known, it is one of the fundamental problems of the quantum theory of solids to obtain the accurate knowledge about the electronic band structure of the actual crystals to which the manifold of physical properties characteristic of solid state can be ascribed at least in principle. Although many efforts along these lines have been made so far and are still being continued, they do not seem to have attained any convincing conclusions even for rather simple crystals.

As for the alkali metals, for which the most thorough investigations have been made so far, the different methods of approximations are known to give rise to some discrepancies in the detailed behaviours of the band structure, although the general behaviours are consistent with each other, getting agreements with experiment, as shown, for sodium, by Lage & Bethe⁽¹⁾ and by Howarth & Jones,⁽²⁾ and, for lithium, by Schiff,⁽³⁾ Kohn & Rostoker,⁽⁴⁾ by Parzen⁽⁵⁾ and by Parmenter.⁽⁶⁾ Even for the simplest metals, therefore, the problem does not seem to be settled completely.

As for the polyvalent metals such as beryllium,⁽⁷⁾ magnesium,⁽⁸⁾ titanium,⁽⁹⁾ iron⁽¹⁰⁾ and nickel,⁽¹¹⁾ the various methods of approximations have been applied to deal with the electronic band structure, however the results of the computation seem to be far less reliable compared with those of alkali metals owing to the polyvalent character of the respective crystals.

Between the alkali metals and the polyvalent metals there are monovalent noble metals, whose band structures have been studied also by many workers. Above all, the most detailed investigations have been performed for the metallic copper by Fuchs,⁽¹²⁾

Krutter,⁽¹³⁾ Tibbs⁽¹⁴⁾ and Howarth^{(15), (16)}, who have adopted the Wigner-Seitz cellular method and the augmented plane wave method, respectively. Although their calculations provide useful knowledge about the band structure of the metallic copper, there still remain discrepancies, qualitative or quantitative, in the detailed behaviours of energy E versus wave number \vec{k} according to the methods of approximations mentioned above. In particular, the final results obtained are remarkably sensitive to the potential to be used, as shown by Howarth.

In view of such unsettled situation in metallic copper, it seems to be highly desirable to study the band structure in copper by the use of the orthogonalized plane wave method (O.P.W.) proposed by Herring⁽¹⁷⁾ and to compare the results with those obtained by other methods, since the O.P.W. method has already been proved to be powerful for alkali metals and even for other non-metallic crystals by Parmenter⁽¹⁸⁾ and Herman.⁽¹⁹⁾ Since the conduction electrons in the monovalent noble metals are known to behave similarly to free electrons on the empirical basis, we may expect that the O.P.W. method gives presumably a good approximation to this problem. In this article, therefore, we shall describe the result of our application of O.P.W. to the metallic copper, which is compared with the previous results.

§ 2. Outline of the method of computation

It is widely recognized that the O.P.W. method has many advantages: at first this method is based upon the secular problem in which the unperturbed wave functions satisfy automatically the boundary conditions and, moreover, the self-consistent procedure can be proceeded, without much labour, in this method. Secondly the potential need not be spherically symmetric about each of the atoms in the crystal. Further, as far as the conduction electrons in metallic copper do concern, O.P.W. expansion will be expected to show presumably good convergence in view of Parmenter's results for the metallic lithium.⁽²⁰⁾

Following Herring's original ideas, each O.P.W. may be written as

$$\chi(\vec{k}+\vec{h}:\vec{r})=V^{-1/2}\exp i(\vec{k}+\vec{h}:\vec{r})-\sum_{\beta}A_{\beta}(\vec{k}+\vec{h})\chi_{\beta}(\vec{k}:\vec{r}), \quad (2.1)$$

where \vec{k} is the reduced wave number vector, \vec{h} a translation vector in the reciprocal lattice, $\chi_{\beta}(\vec{k}:\vec{r})$ the Bloch sum for core-electron and $A_{\beta}(\vec{k}+\vec{h})$ the orthogonalizing coefficient, which is given by

$$A_{\beta}(\vec{k}+\vec{h})=V^{-1/2}(\chi_{\beta}(\vec{k}:\vec{r}), \exp i(\vec{k}+\vec{h}:\vec{r})). \quad (2.2)$$

The variational procedure based on the linear combination of $\chi(\vec{k}+\vec{h}:\vec{r})$ leads to the secular equation, whose matrix elements are given by,

$$(\chi(\vec{k}+\vec{h}:\vec{r}), \chi(\vec{k}+\vec{h'}:\vec{r}))=\delta_{\vec{h},\vec{h'}}-\sum_{\beta}A_{\beta}^{*}(\vec{k}+\vec{h})\cdot A_{\beta}(\vec{k}+\vec{h'}) \quad (2.3)$$

and

$$\begin{aligned} (\chi(\vec{k}+\vec{h}:\vec{r}), H\chi(\vec{k}+\vec{h'}:\vec{r})) &= (\vec{h}+\vec{k})^2\delta_{\vec{h},\vec{h'}}+U[\vec{h'}-\vec{h}] \\ &\quad -\sum_{\beta}E_{\beta}(\vec{k})A_{\beta}^{*}(\vec{k}+\vec{h})A_{\beta}(\vec{k}+\vec{h'}), \end{aligned} \quad (2.4)$$

where $U[\vec{K}] = 1/\Omega_0 \cdot (U(\vec{r}), \exp i(\vec{K}, \vec{r}))_{\Omega_0}$ represents the Fourier coefficient of the crystal potential $U(\vec{r})$.

The core functions

$$\chi_n(\vec{k}; \vec{r}) = N^{-1/2} \sum_v \exp i(\vec{k}, \vec{R}_v) \cdot \varphi_n(\vec{r} - \vec{R}_v) \quad (2.5)$$

are constructed approximately from the atomic orbital $\varphi_n(\vec{r} - \vec{R}_v)$ of Hartree's self-consistent field with exchange for Cu^+ , and further the overlaps between the neighbouring cores are approximately neglected. Then we have $E_n(\vec{k}) = E_{nl}$, which represents atomic energy belonging to the quantum numbers of n and l .

If the potential $U(\vec{r})$ is once given, the relation between the energy E and the wave vector \vec{k} may be given by solving the following secular equation,

$$\det |(\chi(\vec{k} + \vec{h}; \vec{r}), (H - E)\chi(\vec{k} + \vec{h}; \vec{r}))| = 0, \quad (2.6)$$

which is explicitly written by using (2.3) and (2.4).

As pointed out in the literature, the accuracy of the solutions of (2.6) will be mainly governed by the appropriateness of the assumed crystal potential and the number of O.P.W. functions contained in (2.6). The criterion for the latter is assured by checking the good convergence of the expansion in the actual calculations. However, there seems to be no convincing criterion for constructing an appropriate crystal potential, except for the comparison of the final results with the experiment, in spite of a decisively important effect of the crystal potential on the final results.

In fact, Parmenter⁽³⁰⁾ has constructed in his O.P.W. calculation the mentioned crystal potential for Li by placing the neutral Li atom with electron configuration $(1s)^2(2s)^{5/8}(2p)^{3/8}$ at each lattice point, whereas, Herman⁽²⁾ has assumed for each C atom in diamond to be in valence state of $(1s)^2(2s)^1(2p)^{3.5}S$. Although both authors have actually started from the picture of free atoms as mentioned above, we shall attempt to construct the crystal potential in copper by using the model of the free ion cores at lattice points plus free conduction electrons except for a concerned electron. The field of Cu^+ ion at each lattice point is given by the solution of Hartree-Fock equation for Cu^+ ion, i.e.,⁽³¹⁾

$$V_{\text{coul}}(\vec{r}) = -2/r \cdot Z_p(\vec{r}) \quad (2.7)$$

and the contribution from the conduction electrons is approximately replaced by the uniform charge density, following Herring's proposition,⁽³²⁾ by taking account of the fact that Wigner-Seitz method for alkali metals has given actually the nearly uniform charge distribution of conduction electrons except for the very vicinity of the nuclei at lattice points.

In other words, the assumption of the uniform charge density may be taken to represent the starting potential of our procedure. The exchange interaction of a conduction electron with an ion core is taken into account approximately by using Slater's exchange potential, on account of which the potential energy of a conduction electron due to an ion core at each lattice point may be given by

$$v(r) = -2/r \cdot [Z_p(r) + 3(3/32\pi^2)^{1/2} \{\sum_{nl} 2(2l+1)r \cdot P_N^2(nl|r)\}^{1/3}]. \quad (2.8)$$

Then, it follows,

$$U[K] = 3/r_s^3 K \cdot \left[- (2/K) + \int_0^\infty (rv(r) + 2) \sin Kr dr \right],$$

for $K \neq 0$ (r_s = radius of a s -sphere). (2.9)

As shown by Herring,²³⁾ $U[0]$, i.e., the mean value of $U(\vec{r})$ over s -polyhedron, is approximately equal to the mean value averaged over s -sphere of the potential caused by both ion core located at the center and the uniform charge density of other conduction electrons. The mentioned approximation for $U[0]$ seems to be reasonably good in copper, because the first term in (3.2) is essentially equal to $-2/r$ outside the s -sphere and the charge neutrality may be well preserved within the s -sphere.

Thus we obtain,

$$U[0] = (3/r_s^3) \int_0^{r_s} r^2 v(r) dr + (2 \cdot 4/r_s) - 6(9/32\pi^2)^{1/3} \cdot (1/r_s), \quad (2.10)$$

the second and third terms of which denote the contribution from the conduction electrons.

As the ion-core within the crystal is approximately replaced by free Cu^+ ion in the above calculation, our procedure will not be considered to give a satisfactory answer to the problem for the reason that the effect of overlaps or distortions of the relatively large $3d$ -cores is not fully taken into account by such a procedure. Recently, such an effect has actually been partly worked for the electronic band structures of the metallic iron²⁴⁾ but the method adopted there does not seem to be sufficiently satisfactory.

§ 3. Numerical results and discussions

The secular equations of (2.6) are resolved into factors by the use of group theoretical considerations and our actual procedure has been chosen as follows.

For the state corresponding to a fixed \vec{k} -vector, we take into consideration $\chi(\vec{k} + \vec{h}; \vec{r})$ corresponding to the smaller values of $|\vec{k} + \vec{h}|$ in succession, from which the linear combinations are set up to solve the corresponding secular equation. This may be justified from the fact that the mean energy corresponding to each O.P.W., i.e., $\chi(\vec{k} + \vec{h}; \vec{r})$ is approximately determined by the relative magnitudes of $|\vec{k} + \vec{h}|$ and the O.P.W. with lower mean energies are considered to make dominant contributions in the linear combination.

At the edge points in the Brillouin zone, however, not only one wave number vector \vec{k} but also some other vectors $\vec{k} + \vec{h}$ may give the minimum magnitude of $|\vec{k} + \vec{h}|$, as seen in the case of $(2\pi/a)(\frac{3}{4}, \frac{3}{4}, 0)$, $(2\pi/a)(\frac{1}{4}, \frac{1}{4}, 1)$ and $(2\pi/a)(\frac{1}{4}, \frac{3}{4}, 1)$ at a point K in Fig. 1. In such a case we consider only the states which belong to the irreducible representation in the group of $k^{(25)}$ appearing in the linear combinations and the states not

appearing in the linear combinations of the minimum $|\vec{k}+\vec{b}|$ are disregarded, for these correspond certainly to the states of higher energy which are not usually occupied by electrons.

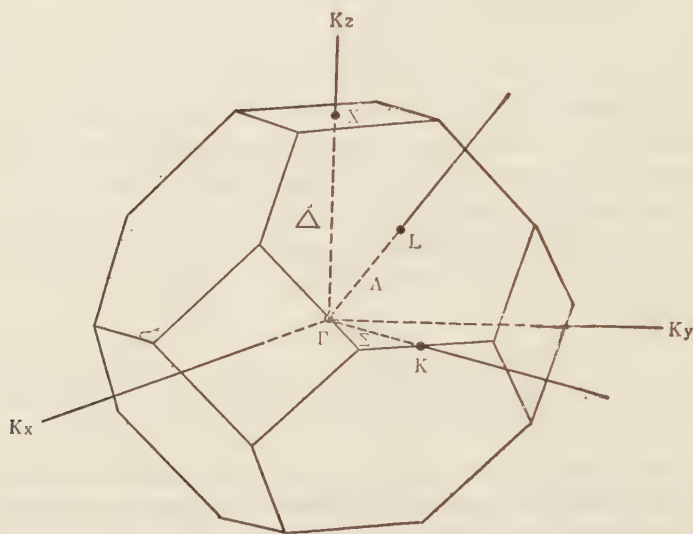


Fig. 1. The first Brillouin zone of a face-centered cubic crystal. Letters indicate points associated with wave functions of particular symmetries.

For the state corresponding to X in Fig. 1 the minimum of $|\vec{k}+\vec{b}|$ becomes $2\pi/a$, $\vec{k}+\vec{b}$ being taken as $(2\pi/a)(0.0.1)$ and $(2\pi/a)(0.0.1)$. Thus, from the group-theoretical procedure, we have X_s and X_p states, of which the eigenfunctions become $\chi(0.0.1:\vec{r}) + \chi(0.0.\bar{1}:\vec{r})$ and $\chi(0.0.1:\vec{r}) - \chi(0.0.\bar{1}:\vec{r})$, respectively.

Actual computations have been carried out along each of the directions in k -space, i.e., $\Delta(0.0.1)$, $\Lambda(1.1.1)$ and $\Sigma(1.1.0)$ in Fig. 1, and in particular at the edge points X, L, K and the origin Γ in the figure, where the considerably rapid convergences are achieved, as shown in Fig. 2. The symmetry along each of the above mentioned directions in k -space becomes lower than those of edge points, so that the reduction of the secular equation in the former cases is incomplete compared with the latter, resulting in the relatively poor convergence. From the comparison with the results on other substances, the degree of the convergence in copper is estimated to be intermediate between those of the alkali metals and the valence crystals as is expected.

In Table I, the values of E computed for a number of k -values are shown.

Since the computed energies have been obtained only for the limited number of k -values, we have attempted to represent $E(\vec{k})$ by the rather simple polynomial of \vec{k} , the coefficients of which are determined by the least square method, as shown in Table II. Since this polynomial-expression is expected to be valid for small k -values, the values of

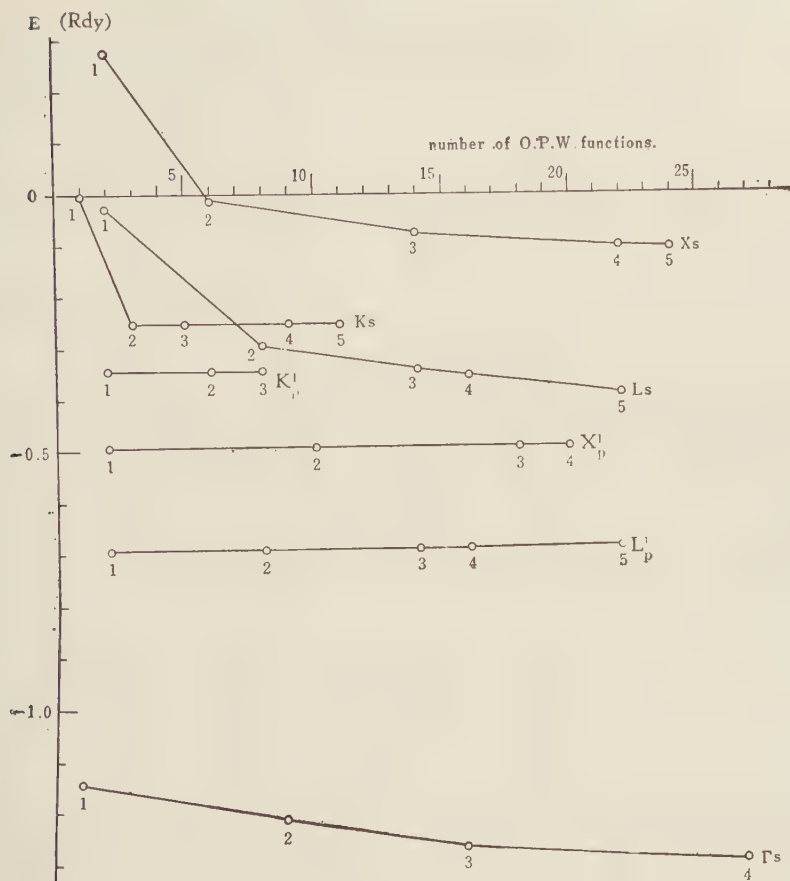


Fig. 2. The degree of convergence in the O.P.W. expansion. The number attached to the points on the curves are the order of the secular equation in the reduced form.

Table I. The energy-values at various wave number k . (Ryd.)

(1.0.0) direction

(1.1.1) direction

(1.1.0) direction

| k | | E | n | k | | E | n | k | | E | n |
|--------|------------|--------|-----|--------|------------|--------|-----|--------|--------------|--------|-----|
| 0 | Γ_s | -1.282 | 27 | 0 | Γ_s | -1.282 | 27 | 0 | Γ_s | -1.282 | 27 |
| 0.2 | A_s | -1.211 | 15 | 0.2 | A_s | -1.204 | 12 | 0.2 | Σ_s | -1.200 | 13 |
| 0.4 | A_s | -1.089 | 15 | 0.4 | A_s | -1.073 | 11 | 0.4 | Σ_s | -1.062 | 9 |
| 0.6 | A_s | -0.882 | 14 | 0.6 | A_s | -0.868 | 11 | 0.6 | Σ_s | -0.850 | 9 |
| 0.9212 | X_p^1 | -0.499 | 20 | 0.7978 | L_p^1 | -0.696 | 24 | 0.7978 | Σ_s | -0.541 | 9 |
| 0.9212 | X_s | -0.108 | 24 | 0.7978 | L_s | -0.393 | 22 | 0.9771 | K_s | -0.260 | 11 |
| | | | | | | | | 0.9771 | K_p^1 | -0.344 | 8 |
| | | | | | | | | 0.7978 | Σ_p^1 | -0.135 | 6 |

n is the number of the O.P.W. functions contained in the expansion.

Table II. The coefficients of the powers of \vec{k} in the $E(\vec{k})$ -curve.

$$E_k = E_0 + E_1 k^1 + E_2 k^1 + 5E_3 (k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2 - 1/5 k^4)$$

| | E_0 | E_1 | E_2 | E_3 |
|-------------|---------|-------|--------|--------|
| For copper | -1.282 | 1.249 | -0.194 | 0.265 |
| For lithium | -0.6832 | 0.723 | 0.039 | -0.033 |

The values for lithium are quoted from the reference 4.

edge points, i.e., X, L and K, have not been used for its computation. Furthermore, the above mentioned expression may be considered not to be sufficiently accurate for the detailed behaviour near the Fermi surface in metallic copper.

The similar energy-expression by a polynomial of \vec{k} has been worked for lithium, by Kohn and Rostoker,²⁶⁾ with the results of $E_2=0.039$, $E_3=-0.033$ which are much nearer to the free electron values than ours for copper. The relatively large values of E_2 and E_3 for the case of copper would be presumably ascribed to the rather large 3d-core of copper ion.

In Figs. 3, 4 and 5 the energy curves along each direction in k -space are shown

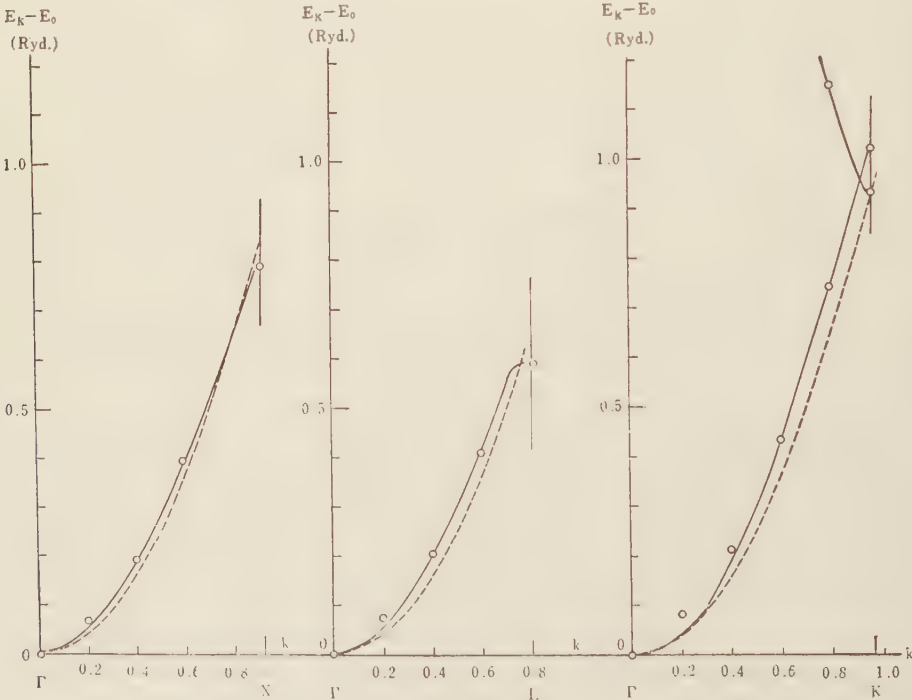


Fig. 3. E versus k along the (0,0,1) direction.

Fig. 4. E versus k along the (1,1,1) direction.

Fig. 5. E versus k along the (1,1,0) direction.

Broken lines express the energy-curves of free electrons.

together with those of the free electrons for comparison.

The lowering of the eigenvalue of the state L_p^1 compared with that of L_s is observed to agree with the suggestion by Mott,²⁷⁾ who pointed out, from the X-ray emission spectrum of copper, that at points near the (1.1.1)-planes inside the first Brillouin zone, the symmetry is predominantly of P -type.

The energy bands of metallic copper have been already computed by the Wigner-Seitz method²⁸⁾ and the augmented plane wave method²⁹⁾ also. In Table III, therefore, the results by the various methods are listed for comparison.

Table III. The results by various methods of approximations.

H-F. ion core potential

(The relative values to Γ_s state, (eV))

| | L_s | L_p^1 | gap | X_s | X_p^1 | gap | K_s | K_p^1 |
|------|-------|---------|-----|-------|---------|-----|-------|---------|
| WS | 8.86 | 10.56 | 1.7 | 11.37 | 10.83 | 0.5 | — | — |
| APW | 7.83 | 9.25 | 1.4 | 9.92 | 11.65 | 1.7 | 6.11 | 5.99 |
| OPW | 12.09 | 7.98 | 4.1 | 15.97 | 10.65 | 5.3 | 13.91 | 12.76 |
| free | 8.66 | | | 11.54 | | | 12.99 | |

Hartree ion core potential

| | L_s | L_p^1 | gap | X_s | X_p^1 | gap | K_s | K_p^1 |
|-----|-------|---------|-----|-------|---------|-----|-------|---------|
| WS | 8.34 | 8.58 | 0.2 | 14.87 | 11.06 | 3.8 | — | — |
| APW | 6.88 | 8.80 | 1.9 | 8.68 | 10.87 | 2.2 | 5.59 | 5.74 |

As is seen in Table III, the energy values of the conduction electrons in the metallic copper given by the O.P.W. method behave in a quite similar way to that of free electrons along each of three directions in k -space, whereas the results by the augmented plane wave method show much deviation from the free electron values along the (1.1.0)-directions. In both directions of (1.0.0) and (1.1.1), however, the band widths obtained by the three methods are nearly equal, while there are some discrepancies in the relative positions of $E(X_s) - E(X_p^1)$ and $E(L_s) - E(L_p^1)$ and in the energy gaps. In fact, our values for the energy gaps are larger than those of the others and are seen to rather resemble to the results of Tibbs.³⁰⁾ These discrepancies might be caused by our use of Slater's exchange potential in the form of free electron approximation, since the contribution from this potential is rather large and then the inaccuracy involved in Slater's averaged exchange potential is taken to have sensitive influence upon the final results.

Although our result does not give an accurate information on the Fermi energy surface, it is inferred from our computation that the magnitude of m/m^* on the Fermi surface will be much smaller in the (1.1.1)-direction.

In view of the results of our computation together with the other data the assumed crystal potential is observed to give rise to remarkably sensitive influence upon the final energy values. It will, therefore, be highly needed to follow the self-consistent procedure in the actual calculation and examine carefully the effect of crystal potential upon the

final result. The investigation of both $4s$ - and $3d$ -bands along these lines are now in progress.

The author should like to express his sincere thanks to Prof. T. Muto and J. Yamashita for suggesting the problem and for their kind guidance throughout the present work.

References

- 1) F. C. von der Lage and H. A. Bethe, *Phys. Rev.* **71** (1947), 612.
- 2) D. J. Howarth and H. Jones, *Proc. Phys. Soc. A* **65** (1952), 355.
- 3) B. Schiff, *Proc. Phys. Soc. A* **67** (1954), 2.
- 4) W. Kohn and Rostoker, *Phys. Rev.* **94** (1954), 1111.
- 5) T. Wainwright and G. Parzen, *Phys. Rev.* **92** (1953), 1129.
- 6) R. H. Parmenter, *Phys. Rev.* **86** (1952), 552.
- 7) C. Herring and A. G. Hill, *Phys. Rev.* **58** (1940), 132.
- 8) S. Raimes, *Phil. Mag.* **41** (1950), 568.
- 9) B. Schiff, *Proc. Phys. Soc. A* **68** (1955), 686.
- 10) M. F. Manning, *Phys. Rev.* **63** (1943), 190; J. B. Greene and M. F. Manning, *Phys. Rev.* **63** (1943), 203; J. Callaway, *Phys. Rev.* **97** (1955), 933; *ibid.* **99** (1953), 500.
- 11) G. C. Fletcher, *Proc. Phys. Soc. A* **65** (1952), 192.
- 12) K. Fuchs, *Proc. Roy. Soc. A* **151** (1935), 585.
- 13) M. Krutter, *Phys. Rev.* **48** (1935), 664.
- 14) S. R. Tibbs, *Proc. Camb. Phil. Soc.* **34** (1938), 89.
- 15) D. J. Howarth, *Proc. Roy. Soc. A* **220** (1953), 513.
- 16) D. J. Howarth, *Phys. Rev.* **99** (1955), 469.
- 17) C. Herring, *Phys. Rev.* **57** (1940), 1169.
- 18) See reference 6.
- 19) F. Herman, *Phys. Rev.* **93** (1954), 1214.
- 20) See reference 6.
- 21) See reference 19.
- 22) D. R. Hartree and W. Hartree, *Proc. Roy. Soc. A* **157** (1936), 490.
- 23) See reference 17.
- 24) J. Callaway, *Phys. Rev.* **99** (1955), 500.
- 25) Boukaert, Smoluchowski and Wigner, *Phys. Rev.* **50** (1936), 58. See also reference 2 and 15. We have used the notations proposed by D. J. Howarth and H. Jones in the present work.
- 26) See reference 4.
- 27) N. F. Mott, *Phil. Mag.* **44** (1953), 187.
- 28) See reference 15.
- 29) See reference 16.
- 30) See reference 14.

Extensions of Variational Methods, III

—Determination of Potential from Phase Shift Function—

Takashi OHMURA (formerly KIKUTA)

Department of Physics, University of Tokyo, Tokyo

(Received April 4, 1956)

The potential $V(r)$ is uniquely constructed from the phase shift function $\eta_l(k)$ in any fixed value l of angular momentum under a certain condition on $V(r)$, if the system has no bound state. The usual variational method is extended so as to give a variation principle for $V(r)$ in that case. We will consider an integral equation in the following form

$$A(k) = \int_0^\infty B(k, r)V(r)dr,$$

where $A(k)$ and $B(k, r)$ are determined by a trial function which is specified by a given $\eta_l(k)$. By means of this equation it is shown that $V(r)$ can be calculated for any $r(\geq 0)$ accurately within errors of the order of Δ^2 , while the errors involved in the trial function are of the order Δ .

When the system has bound states the situation is more complicated, but the phase shift $\eta_l(k)$ uniquely determines the potential for the system as well as the binding energies of all bound states, when we restrict ourselves to a potential of short range which is subjected to the condition,

$$\lim_{r \rightarrow \infty} V(r)e^{2\pi\kappa r} \rightarrow 0, \text{ as } r \rightarrow \infty (\kappa > 0),$$

where $-\kappa$ is the logarithmic derivative for the asymptotic wave function of the lowest level. The arguments given in the case of no bound state remain still valid for such a short range potential. Some generalizations of our method are proposed to include the tensor potential as well as the central one, which will enable us to make use of known phase shifts in the two-nucleon problem.

§ 1. Introduction

We have pointed out in the preceding papers¹⁾ that the usual variational method can be generalized in two different ways. The present paper will be concerned again with another extension of the variational method. To keep the argument simple and clear, we will not enter into a general treatment, but take up a concrete example: Determination of a central potential from the phase shift $\eta_l(k)$ for a fixed angular momentum. This example is of some interest in the theory of scattering and has been investigated by many authors.²⁾⁻¹¹⁾ There exists a complicated problem connected with the bound states, which will be considered in § 3. To avoid unnecessary confusion we will first treat the case in which the system has no bound state, so that the scattering potential $V(r)$ is uniquely determined by the phase shift function $\eta_l(k)$ assuming an appropriate condition on $V(r)$. The usual way of applying the variational method to scattering problems is to find $\eta_l(k)$ for a fixed k from the given $V(r)$. However it seems more reasonable to find $V(r)$ from $\eta_l(k)$, because the knowledges obtainable directly from the analysis of

experiments concern the phase shift rather than the potential $V(r)$. We will formulate a variational method for $V(r)$ in § 2. For two nucleon system the usefulness of this procedure is limited mainly by the following two reasons. The phase shift for a fixed angular momentum (for example, S-states) can not be well determined for all energies owing to the uncertainty of experiments. At very high energy region it seems very probable that the interaction between two nucleons is not represented by a static potential. However, this is a common difficulty so far as we consider the non-relativistic Schrödinger equation, and an attempt of determination of potential is to follow, even if incomplete, the analysis of phase shift. We will mention incidentally the most remarkable point in the method which will be developed in the present paper. In the usual variational methods, the stationary expression for an eigenvalue (or number) enables one to calculate the value of this eigenvalue (or number) accurately with errors only of the order of J^2 , where J indicates the order of errors contained in the trial function. Is it possible to extend this stationary property for an eigenvalue (or number) to the case for a function itself? It is really possible, as discussed in § 2, then we can calculate the values of a function itself accurately with errors of $O(J^2)$ only. The stationary expression for $V(r)$ is shown to be given by (6), and it is proved that the stationary requirement is equivalent to the fundamental Schrödinger equation with specified $\gamma_l(k)$. The closely parallel arguments are given in § 3 for the case when the system has bound states. An application is shown in § 4 for the neutron-proton scattering in S-states, which is described by the following relation,

$$k \cot \gamma_0(k) = -\alpha + (r_0/2)k^2.$$

We see that a simple trial function gives good results. In § 5 two generalizations of the theory are described in order to take into consideration the tensor potential as well as the central one. In the actual phase shift analyses the values of phase shift at discrete energies only can be obtained, and we will give a simple modification in Appendix to use the expression (6) for actual cases.

§ 2. A variation principle for the potential function $V(r)$ in the system without bound state

For simplicity we will take the non-relativistic Schrödinger equation for the S-wave scattering by a central potential $V(r)$:

$$\left\{ \frac{d^2}{dr^2} + k^2 + V(r) \right\} u(k, r) = 0, \quad (1a)$$

$$u(k, 0) = 0, \quad u(k, r) \rightarrow \text{const} \times \sin(kr + \gamma(k)). \quad (1b)$$

where k^2 is the energy of the incident particle in a suitable unit. $u(k, r)$ is real for $k > 0$, $r > 0$. The sign of $V(r)$ is so chosen that an attractive well corresponds to (+). Let the phase shift function $\gamma_l(k)$ be known for all energy values k^2 . N. Levinson²⁾ proved that there exists at most only one potential function $V(r)$ which reproduces the given

phase shift $\eta(k)$, if the system has no bound state and $V(r)$ satisfies the following conditions

$$\int_0^\infty r|V(r)|dr < \infty. \quad (2)$$

$$\int_0^\infty r^2|V(r)|dr < \infty. \quad (3)$$

The conditions (2) and (3) are not trivial, considering that V. Bargmann³⁾ showed two explicit examples of phase equivalent potential even if there is no bound state. We will assume that the given $\eta(k)$ can be reproduced by a certain $V(r)$ which obeys the conditions (2), (3), and consequently $V(r)$ is uniquely determined by $\eta(k)$. Our aim in this section is to give a method to determine $V(r)$ from $\eta(k)$ in a variationally correct way.

Multiplying $u(k, r)$ from the left side of eq. (1), we get

$$-\int_0^\infty u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr = \int_0^\infty u^2(k, r) V(r) dr. \quad (4)$$

We will abbreviate eq. (4) using the following notations,

$$A(k) \equiv -\int_0^\infty u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr, \quad B(k, r) \equiv \{u(k, r)\}^2, \quad (5)$$

with the reduced equation

$$A(k) = \int_0^\infty B(k, r) V(r) dr. \quad (6)$$

Eq. (6) is a Fredholm integral equation of the first kind and $V(r)$ can be determined by this equation if $u(k, r)$ is exactly known. In our case however only the function $\eta(k)$ is known and hence eq. (6) seems useless for determining $V(r)$. However (6) is a fundamental equation in the present variational method as will be shown later.

We first take an arbitrary infinitesimal variation of the wave function $u(k, r)$ as in the usual variational method, which, of course, does not violate the condition (1b). When the wave function $u(k, r)$ is subjected to such variation, $A(k)$ and $B(k, r)$ are correspondingly changed and $V(r)$ also suffer an infinitesimal variation so as to hold eq. (6). Thus we have

$$\delta A(k) = \int_0^\infty \delta B(k, r) V(r) dr + \int_0^\infty B(k, r) \delta V(r) dr, \quad (7)$$

where

$$\delta A(k) \equiv -\delta \int_0^\infty u(k, r) \cdot \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr \quad (8a)$$

$$= -\int_0^\infty \delta u(k, r) \cdot \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr - \int_0^\infty u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) \delta u(k, r) dr,$$

$$\partial B(k, r) \equiv 2u(k, r) \partial u(k, r). \quad (8b)$$

The operator d^2/dr^2 is Hermitian if the condition (1b) is imposed to the wave function, as is verified as follows.

$$\begin{aligned} & \int_0^\infty \partial u(k, r) (d^2 u(k, r) / dr^2) dr \\ &= \partial u(k, r) \cdot du(k, r) / dr \Big|_0^\infty - d(\partial u(k, r)) / dr \cdot u(k, r) \Big|_0^\infty + \int_0^\infty d^2(\partial u(k, r)) / dr^2 \cdot u(k, r) dr. \end{aligned} \quad (9)$$

Using the following properties of ∂u and u ,

$$\partial u(k, 0) = u(k, 0) = 0, \quad \partial u(k, r) \rightarrow \partial c \sin(kr + \eta(k)), \quad u(k, r) \rightarrow c \sin(kr + \eta(k)),$$

we find

$$\partial u(k, r) \cdot du(k, r) / dr \Big|_0^\infty = d(\partial u(k, r)) / dr \cdot u(k, r) \Big|_0^\infty.$$

Hence (9) is reduced to the following equation ;

$$\int_0^\infty \partial u(k, r) \cdot d^2 u(k, r) / dr^2 \cdot dr = \int_0^\infty u(k, r) \cdot d^2(\partial u(k, r)) / dr^2 \cdot dr. \quad (10)$$

Substituting (10) into (8a) and using (1a) and (8b), we have

$$\begin{aligned} \partial A(k) &= -2 \int_0^\infty \partial u(k, r) \cdot \left(-\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr \\ &= 2 \int_0^\infty \partial u(k, r) V(r) u(k, r) dr = \int_0^\infty \partial B(k, r) V(r) dr. \end{aligned} \quad (11)$$

If eq. (11) is compared with eq. (7), it follows that

$$\int_0^\infty B(k, r) \delta V(r) dr = 0. \quad (12)$$

Eq. (12) must hold for all values of $k (\geq 0)$, so we conclude

$$\delta V(r) = 0. \quad (13)$$

Derivation of (13) from (12) is justifiable only if the system has no bound state, as will be discussed in § 3. Eq. (13) means that, if the wave function $u(k, r)$ in $A(k)$ and $B(k, r)$ is replaced by an approximate function $\phi(k, r)$ which differs from $u(k, r)$ by an order of magnitude J , the potential function defined by (6) changes only the order of magnitude J^2 so far as the condition (1b) is not violated. Thus it has been proved that eq. (6) is a stationary expression for $V(r)$.

Next we can prove the converse of this proposition, namely, the requirement that $V(r)$ in eq. (6) is stationary is equivalent to that the wave function $u(k, r)$ must obey eq. (1a) so far as eq. (1b) is maintained. The stationary requirement for $V(r)$ in (6) can be reduced to

$$\delta A(k) = \int_0^\infty \delta B(k, r) V(r) dr, \quad (14)$$

which is seen from (7) and (8). Since the condition (1b) is maintained, eq. (10) is now valid. From (8), (14) and (10) we directly derive the following equation;

$$\int_0^\infty \delta u(k, r) \cdot \left(\frac{d^2}{dr^2} + k^2 + V(r) \right) u(k, r) dr = 0. \quad (15)$$

Eq. (15) must hold for arbitrary variation $\delta u(k, r)$, so we have eq. (1a). This fact suggests that in the present problem there is a variational procedure corresponding to the usual one, however, it is very complicated and is of no use for practical calculation. We merely say that the potential function $V(r)$ is determined by eq. (6) with an accuracy of $O(\mathcal{A}^2)$ if the wave function is chosen with an error of $O(\mathcal{A})$. Next we must solve eq. (6). Such trouble seems inevitable, because we want to determine a function itself rather than an eigenvalue (or a number) as in the usual variational method.

§ 3. The system with bound states

In this section the variation principle given in last section will be extended to the system with bound states. V. Bargmann³⁾ gave explicit examples of phase equivalent potentials which reproduce different binding energies, and moreover showed that there exist continuous families of phase-binding energy equivalent potentials. According to the Gelfand-Levitan theory, the binding energies are completely independent of the phase shift (equivalently the S-matrix). Let the parameter specifying a family of phase-binding energy equivalent potentials be $c (\geq 0)$. We know that there is at most one potential satisfying the conditions (2) ~ (3) to a specified set of $\eta(k)$, κ_i^2 (binding energies, $i=1, \dots, n$) and c_i ($i=1, \dots, n$). Thus $\eta(k)$ is insufficient to determine $V(r)$ completely when the system involves bound states. Nevertheless, in the theory of nucleon-nucleon scattering in the low energy region, we extrapolate the relation $k \cot \eta(k) = -a + (r_0/2)k^2 - \dots$ to the bound state,¹²⁾ so the binding energy of the deuteron is determined by $\eta(k)$ only. Why is this treatment justifiable? The extrapolation is permissible, as discussed by Jost and Kohn⁵⁾, only when the potential satisfies the following restriction,

$$\lim_{r \rightarrow \infty} e^{2\kappa r} V(r) \rightarrow 0, \quad (17)$$

where κ^2 is the largest binding energy of the system, because there is the one-to-one correspondence¹⁰⁾ between zeros of the S-matrix and the binding energies in the region $-\kappa < \text{Im}(k) < 0$. Under the condition (17), all binding energies correspond to zeros of the S-matrix and further there is no redundant zeros. Then the S-matrix (equivalently $\eta(k)$) determines all the binding energies completely.

The arguments developed in § 2 are all valid in the present case too, except for the derivation of (13) from (12). We can find non-vanishing $\delta V(r)$ satisfying (12) as in the paper of Jost and Kohn.⁵⁾ Let $u_i(r)$ be a bound state wave function which satisfy

$$\left(\frac{d^2}{dr^2} - \kappa_i^2 + V(r) \right) u_i(r) = 0, \quad u(0) = 0, \quad u(r) \rightarrow 0 \quad (r \rightarrow \infty)$$

We put $\delta V(r) = 4u_i(r) \cdot du_i(r)/dr \cdot \delta a$, and calculate the left side of (12), because a family of phase-binding energy equivalent potential is obtained by integration of this infinitesimal increment.

$$\begin{aligned} 4\delta a_i \int_0^\infty u_i^2 u_i' dr &= 2 \left\{ \int_0^\infty u_i u (u_i' u + u_i u') dr + \int_0^\infty u_i u (u_i' u - u_i u') dr \right\} \delta a_i \\ &= 2\delta a_i \int_0^\infty u_i u (u_i u)' dr - (k^2 + \kappa_i^2) \left\{ \int_0^\infty u_i(r) u(r) dr \int_0^\infty u_i(r') u(r') dr' + \int_0^\infty u_i(r) u(r) dr \right. \\ &\quad \times \left. \int_0^r u_i(r') u(r') dr' \right\} \delta a_i \\ &= \delta a_i (u_i u)^2 \Big|_0^\infty - (k^2 + \kappa_i^2) \left\{ \int_0^\infty u_i u dr \right\}^2 \delta a_i = 0, \end{aligned}$$

where we have used the following relations

$$\begin{aligned} u_i(r) u'(r) - u(r) u_i'(r) &= (k^2 + k_i^2) \int_r^\infty u_i(r') u(r') dr', \\ \int_0^\infty u_i(r) u(r) dr \int_r^\infty u_i(r') u(r') dr' &= \int_0^\infty u_i(r) u(r) dr \int_0^r u_i(r') u(r') dr'. \end{aligned}$$

Thus (13) does not follow (12). However if we set the requirement (17) on $V(r)$, the above $\delta V(r)$ is not permissible, because $\delta V(r)$ has the following asymptotic form for large r ,

$$\delta V(r) = 4u_i(r) u_i'(r) \rightarrow \text{const} \times e^{-2\kappa_i r},$$

and consequently $e^{2\kappa_i r} \cdot \delta V(r) \not\rightarrow 0$ ($r \rightarrow \infty$). Thus the condition (17) again gives an effective criterion for choosing a special potential from a family of phase-binding energy equivalent potential. Thus we have found that the arguments given in § 2 are all valid if we know a priori that that potential satisfies (17).

§ 4. An example: $k \cot \eta = -\alpha + (r_0/2)k^2$.

In the present section we will apply the variational method for $V(r)$ so far developed in the preceding sections to neutron-proton scattering assuming that the scattering is described by a central static potential. As an example, let the S-wave phase shift be given in the following form

$$k \cot \eta(k) = -\alpha + (r_0/2)k^2, \quad (18)$$

which has been used to an approximate description of neutron-proton scattering in the low energy region¹²⁾. We assume the following form for the trial function,

$$\phi(k, r) = (1 - e^{-cr}) \cos kr + \cot \eta(k) \sin kr \quad (19)$$

where c is an adjustable parameter. $A(k)$ and $B(k, r)$ in eq. (6) are given by

$$A(k) = c \{ 3/4 - (c^2 + 2k^2) / (4c^2 + 4k^2) \} + k \cot \eta(k),$$

$$B(k, r) = \phi^2(k, r).$$

Eq. (6) can be solved analytically only for the function $\eta(k)$, and in general we must take a numerical solution. Since the integral should be extended to infinity, we must adopt a suitable approximation to simplify the numerical solution. Here it is assumed that $V(r)$ can be expanded in a Dirichlet series

$$V(r) = \sum_{j=1}^{\infty} a_j e^{-\lambda_j r}, \quad (20)$$

We assume further that $V(r)$ is well approximated using only finite terms (say n -terms) of (20). Then Eq. (6) is reduced in a simultaneous linear equations of order n ,

$$A(k_i) = \sum_{j=1}^{\infty} a_j \int_0^{\infty} e^{-\lambda_j r} B(k_i, r) dr, \quad (i=1, \dots, n). \quad (21)$$

If λ_j are chosen suitably and n is increased, the solution of (21)

$$V(r) = \sum_{j=1}^n a_j e^{-\lambda_j r} \quad (22)$$

will converge to the exact solution (20) with fair accuracy. But eq. (21) can be interpreted in another way. In an actual analysis of experiments we know the values of phase shift for only discrete values of energy (let this number be n). In this case it seems natural to assume the form of $V(r)$ as (22) with the help of meson theory: $\lambda_j = j/\mu$, μ : Compton wave length of meson. Then we can prove that (21) provides us a stationary expression for $a_j (j=1, \dots, n)$. This proof will be given in Appendix.

The exact solution for (18) is expressed in an analytical form¹³⁾ as

$$V(r) = \frac{2\nu(1-\nu)l^2 e^{-lr}}{[1-\nu(1-e^{-lr})]^2} = \sum_{j=1}^{\infty} d_j e^{-jlr}, \quad \nu = (l-b)/2l,$$

$$l = 2(1 + \sqrt{1-2\alpha r_0})/r_0, \quad b = 2(-1 + \sqrt{1-2\alpha r_0})/r_0. \quad (23)$$

The scattering length $a(=1/\alpha)$ and the effective range r_0 in the neutron-proton system are

$$a = 5.388 \times 10^{-13} \text{ cm}, \quad r_0 = 1.70 \times 10^{-13} \text{ cm}, \quad \alpha r_0 = 0.3155: \quad \text{for the spin triplet state.}$$

r_0 is taken to be the unit of length hereafter. Then the decreasing factor l in (23) are $l=3.215$ (for triplet). According to the meson theory the asymptotic behavior of nuclear potential has the form

$$V(r) \rightarrow \text{const } e^{-1.189r}/r,$$

where the meson mass is assumed to be of 270 electron mass. We will take the following values as λ_j

$$\lambda_j = 3, 6, 9, 12.$$

These λ_j all satisfy the condition (17). Table I shows the results obtained in several approximations using only three terms of (22). ($n=3$). ϵ involved in the trial function is 2. $k^2=1$ corresponds to the incident neutron energy in lab. system 28.7 Mev. We

| r | 0.0 | 0.2 | 0.5 | 1.0 | 2.0 |
|------------------|--------|--------|--------|--------|--------|
| Approx. (1) | 1.756 | 5.863 | 4.003 | 1.099 | 0.058 |
| " (2) | -4.831 | 4.605 | 4.108 | 1.191 | 0.063 |
| " (3) | 28.766 | 4.518 | 3.883 | 1.298 | 0.071 |
| Correct solution | 4.8600 | 5.1411 | 3.8557 | 1.2024 | 0.0549 |

Table I. The values of $V(r)$ obtained from (21) and (22) are compared with the exact solution.

Approx. (1): $k_i=0, 0.5, 1$ and $\lambda_j=3, 6, 9$

Approx. (2): $k_i=0, 1, 3$ and $\lambda_j=3, 6, 9$

Approx. (3): $k_i=0, 2, 4$ and $\lambda_j=3, 6, 12$.

see from Table I that three approximations have comparable accuracy and that the potential $V(r)$ is not well determined in the vicinity of the origin. This is expected from eq. (6), because $B(k, r)$ behaves $\text{const} \times r^2$ for small r . But mathematically we can determine the potential accurately in the vicinity of the origin using the very high energy data, if we assume that the interaction is well represented by a static potential in that region.

§ 5. To include the tensor potential

To apply the present method to nucleon-nucleon scattering, both the tensor force should be taken into consideration as well as the Coulomb force. Then we must get a variation principle for both potentials. In the actual phase shift analyses, we can not get the values of phase shifts of several waves for all energies. But at first we assume that this is possible, and will generalize the method to include the tensor potential as well as the central one. A simple modification which is necessary in actual analysis will be given in Appendix.

The arguments developed below are closely parallel those in § 2. The Coulomb force will be neglected to avoid unnecessary confusion, though it is trivial to include Coulomb interaction. Since the equations in the spin singlet states are essentially the same as in the case of central forces only, we may treat the triplet case where two states with orbital angular momenta $J-1$ and $J+1$ (J : total angular momentum of the system) couple each other,¹⁴⁾

$$\left\{ \frac{d^2}{dr^2} + k^2 - \frac{(J-1)J}{r^2} + \left[V(r) - \frac{2(J-1)}{(2J+1)} W(r) \right] \right\} u(r) = - \frac{6\sqrt{J(J+1)}}{(2J+1)} W(r) w(r), \quad (24a)$$

$$\left\{ \frac{d^2}{dr^2} + k^2 - \frac{(J+1)(J+2)}{r^2} + \left[V(r) - \frac{2(J+2)}{(2J+1)} W(r) \right] \right\} w(r) = - \frac{6\sqrt{J(J+1)}}{(2J+1)} \times W(r) u(r). \quad (24b)$$

These equations (24) may be abbreviated using the following matrix representation.

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} -d^2/dr^2 - k^2 + (J-1)J/r^2, & 0 \\ 0, & -d^2/dr^2 - k^2 + (J+1)(J+2)/r^2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1, & 0 \\ 0, & 1 \end{bmatrix} \\ \mathbf{C} &= \begin{bmatrix} -2(J-1)/(2J+1), & 6\sqrt{J(J+1)}/(2J+1) \\ 6\sqrt{J(J+1)}/(2J+1), & -2(J+2)/(2J+1) \end{bmatrix} \quad \mathbf{v}(r) = \begin{bmatrix} u(r) \\ w(r) \end{bmatrix}. \end{aligned}$$

Eqs. (24) are reduced to

$$\mathbf{A}\mathbf{v}(r) = \{\mathbf{B}V(r) + \mathbf{C}W(r)\} \mathbf{v}(r), \quad (25)$$

(24) or (25) have two independent solutions satisfying $u(0) = w(0) = 0$, which will be chosen in the following way,

$$\mathbf{v}_\alpha(k, r) = \begin{cases} u_\alpha(r) \rightarrow \cos \varepsilon(k) \cdot \sin \{kr - ((J-1)\pi/2) + \eta_\alpha(k)\}, \\ w_\alpha(r) \rightarrow \sin \varepsilon(k) \cdot \sin \{kr - ((J+1)\pi/2) + \eta_\alpha(k)\}, \end{cases} \quad (26a)$$

$$(26b)$$

$$\mathbf{v}_\beta(k, r) = \begin{cases} u_\beta(r) \rightarrow -\sin \varepsilon(k) \cdot \sin \{kr - ((J-1)\pi/2) + \eta_\beta(k)\}, \\ w_\beta(r) \rightarrow \cos \varepsilon(k) \cdot \sin \{kr - ((J+1)\pi/2) + \eta_\beta(k)\}, \end{cases} \quad (26c)$$

$$(26d)$$

where η_α and η_β are the eigenphase shifts of the 2×2 S -matrix for the state J , and ε is a mixing parameter. η_α , η_β and ε represent three-real and independent parameters which characterize the S -matrix completely. Multiplying \mathbf{v}_α from the left side of (25); $\mathbf{A}\mathbf{v}_\alpha = \{\mathbf{B}V + \mathbf{C}W\} \mathbf{v}_\alpha$, we have

$$A_\alpha(k) = \int_0^\infty \{B_\alpha(k, r)V(r) + C_\alpha(k, r)W(r)\} dr, \quad (27a)$$

Similarly,

$$A_\beta(k) = \int_0^\infty \{B_\beta(k, r)V(r) + C_\beta(k, r)W(r)\} dr, \quad (27b)$$

where

$$\begin{aligned} A_\alpha(k) &\equiv - \int_0^\infty \left(u_\alpha \cdot \frac{d^2 u_\alpha}{dr^2} + w_\alpha \cdot \frac{d^2 w_\alpha}{dr^2} \right) dr - k^2 \int_0^\infty (u_\alpha^2 + w_\alpha^2) dr \\ &\quad + \int_0^\infty \{ (J-1)J \cdot u_\alpha^2/r^2 + (J+1)(J+2) w_\alpha^2/r^2 \} dr, \\ B_\alpha(k, r) &\equiv u_\alpha^2 + w_\alpha^2, \dots \dots \dots \\ \dots \dots C_\beta(k, r) &\equiv \left\{ - \frac{2(J-1)}{2J+1} \right\} u_\beta^2 + \frac{12\sqrt{J(J+1)}}{2J+1} u_\beta w_\beta + \left\{ \frac{-2(J+2)}{2J+1} \right\} w_\beta^2. \end{aligned} \quad (27c)$$

We can prove that (27) is a stationary expression for both $V(r)$ and $W(r)$, when the phase shift functions $\gamma_\alpha(k)$ and $\gamma_\beta(k)$ in (26) are specified. Taking variations $\delta v(r)$, the formulas corresponding to (7) and (8) follow,

$$\delta A_\alpha(k) = \int_0^\infty \delta_\alpha B_\alpha(k, r) V(r) dr + \int_0^\infty \delta C_\alpha(k, r) W(r) dr + \int_0^\infty B_\alpha(k, r) \delta V(r) dr + \int_0^\infty C_\alpha(k, r) \delta W(r) dr, \quad (28)$$

and the corresponding one for β . The operator d^2/dr^2 is again Hermitian if $\gamma_\alpha(k)$ and $\gamma_\beta(k)$ are specified. Namely,

$$\int_0^\infty \partial u_\alpha \cdot \frac{d^2 u_\alpha}{dr^2} dr = \int_0^\infty u_\alpha \frac{d^2 (\partial u_\alpha)}{dr^2} dr, \quad \int_0^\infty \partial w_\alpha \cdot \frac{d^2 w_\alpha}{dr^2} dr = \int_0^\infty w_\alpha \frac{d^2 (\partial w_\alpha)}{dr^2} dr.$$

Then we find that

$$\begin{aligned} \delta A_\alpha(k) &= -2 \int_0^\infty \left\{ \partial u_\alpha \cdot \frac{d^2 u_\alpha}{dr^2} + \partial w_\alpha \cdot \frac{d^2 w_\alpha}{dr^2} + k^2 (\partial u_\alpha \cdot u_\alpha + \partial w_\alpha \cdot w_\alpha) \right. \\ &\quad \left. + \frac{(J-1)J}{r^2} \partial u_\alpha \cdot u_\alpha + \frac{(J+1)(J+2)}{r^2} \partial w_\alpha \cdot w_\alpha \right\} dr, \\ &= 2 \int_0^\infty (\partial u_\alpha \cdot u_\alpha + \partial w_\alpha \cdot w_\alpha) V dr + 2 \int_0^\infty \left\{ [-2(J-1)/(2J+1)] \partial u_\alpha \cdot u_\alpha \right. \\ &\quad \left. + \frac{6\sqrt{J(J+1)}}{2J+1} \partial u_\alpha \cdot w_\alpha + \frac{6\sqrt{J(J+1)}}{2J+1} \partial w_\alpha \cdot u_\alpha - \frac{2(J+2)}{2J+1} \partial w_\alpha \cdot w_\alpha \right\} W dr. \\ &= \int_0^\infty \delta B_\alpha(k, r) V(r) dr + \int_0^\infty \delta C_\alpha(k, r) W(r) dr, \end{aligned}$$

where use has been made of eqs. (24). Eq. (28) is now reduced to

$$\int_0^\infty B_\alpha(k, r) \delta V(r) dr + \int_0^\infty C_\alpha(k, r) \delta W(r) dr = 0. \quad (29a)$$

Similarly we also get

$$\int_0^\infty B_\beta(k, r) \delta V(r) dr + \int_0^\infty C_\beta(k, r) \delta W(r) dr = 0. \quad (29b)$$

From (29a) and (29b) we can get the relation

$$\delta V(r) = 0, \quad \delta W(r) = 0.$$

if there is no bound state. When the bound states exist, as in the n - p system, there may be a condition corresponding to (17) under which the potentials $V(r)$ and $W(r)$ are uniquely determined from (29). This condition will be satisfied in the two nucleon system. We should notice that $\varepsilon(k)$ appeared in (26) is considered as an variational

* This fact will be suggested by eq. (9) in the paper: T. Kikuta, Prog. Theor. Phys. 12 (1954), 10.

parameter, hence we need not know $\varepsilon(k)$ with accuracy as $\eta(k)$.

Next we will consider some different case. The α and β -waves behave as if the waves are in $l=J-1$ and $l=J+1$ respectively. As far as the low energy region is concerned, the α -wave gives more contribution to the cross section than the β -wave does. If $\eta_\alpha(k)$ and $\varepsilon(k)$ are known for all energies rather than $\eta_\beta(k)$, we can construct other variation principle for $V(r)$ and $W(r)$. The arguments followed are similar to (26)–(29), therefore we will only give the results. The stationary expressions for $V(r)$ and $W(r)$ are now given by*

$$\begin{aligned} A_\alpha(k) &= \int_0^\infty \{B_\alpha(k, r)V(r) + C_\alpha(k, r)W(r)\} dr, \\ A_\tau(k) &= \int_0^\infty \{B_\tau(k, r)V(r) + C_\tau(k, r)W(r)\} dr, \end{aligned} \quad (30)$$

where A_α , B_α and C_α are the same as (27a), and the quantities with γ are related to both α and β . Expressing them in a simple way, we have

$$A_\alpha(k) \equiv \int_0^\infty \mathbf{v}_\alpha \cdot \mathbf{A} \mathbf{v}_\alpha dr, \quad B_\alpha(k, r) \equiv \mathbf{v}_\alpha \cdot \mathbf{B} \mathbf{v}_\alpha, \quad \dots \quad (27c)$$

$$A_\tau(k) \equiv \int_0^\infty \mathbf{v}_\beta \cdot \mathbf{A} \mathbf{v}_\alpha dr, \quad B_\tau(k, r) \equiv \mathbf{v}_\beta \cdot \mathbf{B} \mathbf{v}_\alpha, \quad C_\alpha(k, r) \equiv \mathbf{v}_\beta \cdot \mathbf{C} \mathbf{v}_\alpha. \quad (31)$$

$\eta_\beta(k)$ involved in \mathbf{v}_β of (31) is a variational parameter, on the other hand the trial function for (30) must have correct $\eta_\alpha(k)$ and $\varepsilon(k)$.

There are several points to be taken into consideration for actual application to two nucleon system, for which we will discuss in Appendix.

§ 6. Concluding remarks

In the usual variational methods, an eigenvalue (or number) can be calculated with accuracy $O(\mathcal{J}^2)$ by using the stationary expression of this eigenvalue (or number), where \mathcal{J} is a quantity which gives measure of errors contained in the trial function. We have extended this advantageous point of the variational methods to the problem in which the function $V(r)$ itself is to be determined from the phase shift function $\eta_l(k)$. The stationary expression of the function $V(r)$ itself has been found to be eq. (6), and by using (6) we can calculate the value of $V(r)$ with accuracy $O(\mathcal{J}^2)$ for all $r(\geq 0)$. A new aspect developed in the present paper concerns the fact that we determine the function itself instead of an eigenvalue (or a number).

Finally we will mention other methods for constructing $V(r)$ from $\eta_l(k)$. The fundamental formula for determining the potential $V(r)$ from a phase shift function $\eta_l(k)$ without using the wave function $u(k, r)$ is the Gelfand-Levitan¹⁾ equation which is derived in Jost and Kohn's paper⁵⁾. We must determine a "spectral function" by $\eta_l(k)$, and construct a function $g(p, s)$ from the spectral function. A linear integral equation of the following type should be solved for all values of a parameter p ,

$$K(p, s) + g(p, s) + \int_0^p g(s, t) K(r, t) dt = 0. \quad (16)$$

The potential $V(r)$ is given by $dK(p, p)/dp$. This theory is very interesting in the theory of the S-matrix. Jost and Kohn^(7,8) have also developed a constructing method of potential in the form of a series. They are useful for a special function $\gamma(k)$ which corresponds to the Eckert potential for which eq. (1) is solved analytically, but it seems extremely difficult to apply to the general type of $\gamma(k)$. On the other hand eq. (6) is solvable approximately for general cases with minimum effort.

The author would like to express his sincere thanks to Professor T. Yamanouchi for his discussions and many corrections in the manuscript of this paper.

Appendix

Let the potential $V(r)$ be expressed as (22) and the values of phase shift at n discrete energy values be known. Then we can determine $a_j (j=1, 2, \dots, n)$ from (21). In this appendix, we shall show that eqs. (21) are stationary expressions for $a_j (j=1, 2, \dots, n)$. Taking a variation in both sides of (21), we have

$$\delta A(k_i) = \sum_{j=1}^n a_j \int_0^\infty e^{-\lambda_j r} \delta B(k_i, r) dr + \sum_{j=1}^n \delta a_j \int_0^\infty e^{-\lambda_j r} B(k_i, r) dr. \quad (i=1, \dots, n) \quad (A1).$$

The left hand side of (A1) is reduced as follows.

$$\begin{aligned} \delta A(k) &= - \int_0^\infty \delta u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr - \int_0^\infty u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) \delta u(k, r) dr, \\ &= -2 \int_0^\infty \delta u(k, r) \left(\frac{d^2}{dr^2} + k^2 \right) u(k, r) dr, \\ &= 2 \int_0^\infty \delta u(k, r) V(r) u(k, r) dr, \\ &= \sum_{j=1}^n a_j \int_0^\infty e^{-\lambda_j r} \delta B(k, r) dr, \end{aligned} \quad (A2)$$

where the use has been made of eq. (10). Comparing (A.2) with (A.1), it follows that

$$\sum_j B_{ij} \delta a_j = 0, \quad (i=1, \dots, n), \quad B_{ij} \equiv \int_0^\infty e^{-\lambda_j r} B(k_i, r) dr.$$

Then we conclude that $\delta a_j = 0$ ($j=1, \dots, n$), provided that the determinant: $|B_{ij}|$ does not vanish. The extension to the case where the interaction potential has a non-central part is straightforward. Let the phase shifts γ_α and γ_β (or γ_α and ϵ) be known at n energy values from the analyses of experiments. We again assume, as in § 4, that $V(r)$ and $W(r)$ are a linear combination of several suitable functions,

$$V(r) = \sum_{i=1}^m a_i f_i(r); \quad \text{central potential}$$

$$W(r) = \sum_{i=m+1}^{2n} a_i f_i(r) ; \quad \text{tensor potential.}$$

Eqs. (23) are now reduced to simultaneous and inhomogeneous linear equations of order $2n$, where the unknown variables are $a_i (i=1, 2, \dots, 2n)$. This set of equations can be proved to give stationary expressions for a_i . The method of proof is quite similar in the case of central potential.

The nuclear forces are represented by the static potentials approximately in some moderate energy region, for example $0 \sim 300$ Mev in the incident nucleon energy. In this energy region we can obtain several knowledges on the values of phase shift analysis. From these knowledges for the phase shifts, we can construct the interaction potentials except the vicinity of the origin using the variational method as discussed in this paper in detail.

References

- 1) T. Kikuta, Prog. Theore. Phys. **14**(1955), 457; **15** (1956), 50.
- 2) N. Levinson, Kgl. Danske Videnskab Selskab, Mat.-fys. Medd. **25**, No. 9 (1949).
- 3) V. Bargmann, Phys. Rev. **75**(1949), 301; Rev. Mod. Phys. **21**(1949), 488.
- 4) I. M. Gelfand and B. M. Levitan, Doklady Akad. Nauk. S. S. S. R. n Ser. **77**, 557 (1951); Izvestiya Akad. Nauk. S. S. S. R. **15** (1951) 309.
- 5) R. Jost and W. Kohn, Phys. Rev. **88** (1952), 382; Kgl. Danske Videnskab Selskab, Mat.-fys. Medd. **27** No. 9 (1953).
- 6) B. Holmberg, Nuovo Cimento **9** (1952), 597.
- 7) C. E. Fröberg, Arkiv Mat. Astron. Fys. **34A** No. 28(1948); **36A** No. 11 (1949); Arkiv Fys. **3**, No. 1 (1951).
- 8) R. Jost and W. Kohn, Phys. Rev. **87**(1952), 977.
- 9) S. T. Ma, Phys. Rev. **69**(1946) 668; **71** (1947); 159, 210.
- 10) R. Jost, Helv. Phys. Acta. **20** (1947), 256.
- 11) N. G. van Kampen, Phil. Mag. **42** (1951), 851; K. Wildermuth, Zeit. f. Physik, **127** (1950), 85; N. Hu, Phys. Rev. **74** (1948), 131.
- 12) J. M. Blatt and J. D. Jackson, Phys. Rev. **76** (1949), 38.
- 13) For example, see reference 10).
- 14) For example, J. M. Blatt and L. C. Biedenharn, Phys. Rev. **86** (1952), 399.
- 15) For example, G. Breit, *Proceeding to the 4th annual Rochester Conference of High Energy Nuclear Physics* (1954), p 19.
A. Garren, Phys. Rev. **101** (1956), 419.
R. M. Thaler and J. Bengtson, **94** (1954), 679.
C. A. Klein, Nuovo Cimento **I**(1955), 581; **II** (1955), 38.

Letters to the Editor

State Vectors in the Collective Description

Toshiyuki Nishiyama

*Department of Physics, Osaka University,
Osaka*

July 2, 1956

In the theory of collective description proposed by Bohm¹⁾ as well as in those of others^{2),3)}, there appears a set of subsidiary conditions. If these conditions were ignored, some fictitious states would appear in the solution of the eigenvalue problem. In order to take account of these conditions, we propose a method to transform a subsidiary condition into another one for which the state vectors are more easily obtained. Our method will be effective in dealing with such cases that the collective motion becomes a harmonic oscillation.

We may consider either a many fermion system or a many boson system. Here it is assumed that a function of n particle coordinates, say $\xi = \xi(x_1, x_2, \dots, x_n)$, which will become the coordinate of the collective oscillation, has its canonical conjugate π determined by

$$\pi = i/\hbar \cdot [T, \xi], \quad [\xi, \pi] = i\hbar, \quad (1)$$

where T is the kinetic energy of particles. In the scheme of second quantization the corresponding canonical quantities ξ and π are given by functionals of the quantized wave functions. A typical example for ξ is the center-of-mass coordinate of a system of particles moving in a parabolic potential. As pointed out by Tomonaga¹⁾ and others^{2),3)}

the coordinate of the plasma oscillation in a metal or that of the surface oscillation of a nucleus satisfies the relation (1) in some approximate meaning. We have proposed a method in the previous paper²⁾ to describe the plasma oscillation by an auxiliary field. The same procedure has been made independently by Miyazima³⁾ for the surface oscillation of a nucleus. In compensation for introducing the redundant degrees of freedom of such an auxiliary field, a set of subsidiary conditions has been imposed on the state vectors as in Bohm's collective description.

Introducing an auxiliary variable Q and its canonical conjugate P , we can go over to the collective representation by a contact transformation of the form

$$\left. \begin{aligned} U &= U_1 U_2 U_1, \\ U_1 &= \exp(i\xi P/\hbar), \\ U_2 &= \exp(-i\pi Q/\hbar), \\ [Q, P] &= i\hbar, \end{aligned} \right\} \quad (2)$$

The subsidiary condition imposed on the state vectors becomes

$$\xi \Psi = 0. \quad (3)$$

In this collective representation the Hamiltonian becomes commutative with both ξ and π ²⁾. This implies that the Hamiltonian is given in terms of a set of canonical quantities independent of ξ and π . The condition (3) may be expressed as

$$\partial \Psi(\eta_1, \eta_2, \dots; P) / \partial \pi = 0, \quad (4)$$

by which it is implied that the state vectors are expressed by functionals of internal quantities; η_1, η_2, \dots , commutative with ξ . In the case of the center-of-mass

motion the internal variables orthogonal to the center-of-mass momentum are obtained by an orthogonal transformation in the configuration space with respect to momenta of particles. In this case, any state vector does not involve the center-of-mass momentum π_c so that (3) is valid for the center-of-mass coordinate ξ_c . For general cases, Lipkin et al.⁵⁾ pointed out that a state vector multiplied by the δ -function $\delta(\xi)$ satisfies the subsidiary condition of type (3). From our point of view, however, it is the more precise interpretation of (3) that the states are independent of π than that the states vectors are multiplied by the δ -function.

Now, we add a Hamiltonian of a harmonic oscillation given by

$$H_h = \pi^2/2 + \Omega^2 \xi^2/2, \quad (5)$$

to the total Hamiltonian H expressed in the collective representation. The frequency Ω is adjusted so that H_0 is simplified as far as possible. In most cases such that the collective motion becomes a harmonic oscillation, the modified Hamiltonian $H_0 = H + H_h$ has a simpler form than H . We consider the eigenvalue problem of H in the representation that the Hamiltonian H_0 is diagonal. Then every eigenstate of H is given by a linear combination of eigenstates of H_0 . In general, such a linear combination does not satisfy the subsidiary condition (3). If the eigenvalue problem of H were solved directly, it would always be possible to make every eigenstate satisfy (3), because H is written in terms of a set of canonical quantities independent of ξ and π . The reason why there appear some eigenstates of H that do not satisfy (3) is that such states are multiplied by some functionals of π . Here, introducing

the destruction operator defined by

$$A = \pi - i\Omega\xi, \quad (6)$$

we classify the eigenstates of H given by linear combinations of the eigenstates of H_0 . The eigenstates satisfying the condition

$$A\Psi = 0, \quad (7)$$

are called proper states and of the other eigenstates, those satisfying (3) are called true states, and the remaining are called improper states. Concerning the true states we get the following theorem:

Theorem

An eigenstate of H becomes a true state, if and only if it takes the form

$$\Psi = \exp(\pi^2/2\hbar\Omega)\Psi_p, \quad (8)$$

where Ψ_p is a proper state.

Proof

Using the following identity

$$\xi \exp(\pi^2/2\hbar\Omega) = i \exp(\pi^2/2\hbar\Omega) A/\Omega, \quad (9)$$

and (7), we can see that a state expressed by (8) becomes an eigenstate of H satisfying (3). In general, an eigenstate of H is given by a functional of the internal quantities η_1, η_2, \dots and of the auxiliary variable P conjugate to Q multiplied by a functional of π . An improper state is multiplied by an arbitrary functional of π , while a proper state is multiplied by the Gaussian functional; $\exp(-\pi^2/2\hbar\Omega)$, for which we get $A \exp(-\pi^2/2\hbar\Omega) = 0$. Every true state is different from a proper state only by this Gaussian functional and given by (8).

The above method has been applied to the well-known problem of coupled harmonic oscillators⁶⁾ and the many boson problems.⁷⁾

A fuller presentation will be given in the following paper.

- 1) D. Bohm and D. Pines, *Phys. Rev.* **92** (1953), 609.
- 2) T. Nishiyama, *Prog. Theor. Phys.* **14** (1955), 37.
- 3) T. Miyazima and T. Tamura, *Prog. Theor. Phys.* **15** (1956), 255; T. Marumori and E. Yamada, *ibid.* **13** (1955), 557 (L); Y. Watanabe, *ibid.* **16** (1956), 1.
- 4) S. Tomonaga, *Prog. Theor. Phys.* **5** (1950), 544; **13** (1955), 467, 482.
- 5) H. J. Lipkin, A. de Shalit and I. Talmi, *Nuovo Cimento* **2** (1955), 773.
- 6) For example, H. R. Post, *Proc. Phys. Soc.* **A66** (1953), 649 (L).
- 7) N. Bogolubov, *J. Phys.* **11** (1947), 23; T. Nishiyama, *Prog. Theor. Phys.* **12** (1954), 265.

Some Remarks on the Nuclear Level Spacing

Yoichi Fujimoto, Fumio Fukuzawa,
and Sueji Okai

*Department of Physics, Kyoto University,
Kyoto*

July 5, 1956

Because a heavy nucleus has many degrees of freedom, at sufficiently high excitation energies the nuclear states become too numerous to treat individually and only a statistical description is possible. Weisskopf¹⁾ has derived the semi-empirical nuclear level density formula based on the statistical model of nucleus. It should be noted that the level spacings obtained from Weisskopf's formula are only those averaged in the small energy region of a given excitation energy.

Recently, the parameters of nuclear energy levels at excitation energy just above neutron binding have been studied for heavy nuclides with improved resolution. Various regularities have been observed^{2),3)} with regard to the widths and spacings of energy levels. In the work of Harvey et al.,²⁾ it was found that within experimental error the reduced neutron widths had an exponential distribution in size for each nuclide that was investigated.

For level spacing (D), on the other hand, the similar distribution may be expected to hold. This can be inferred from the fact⁴⁾ that if the sum of the reduced neutron widths *vs* neutron energy is plotted, it becomes a straight line within experimental error. Therefore, the first aim in this note is to point out that level spacing has also an exponential distribution in size. When the number of level spacings larger than a given value D were plotted as a function of D in a given energy range, the resulting distribution was found to be consistent with an exponential distribution, implying that the differential distribution, that is, the number of level spacings as a function of their size, was also exponential.

In order to investigate this distribution with better statistical accuracy, the combination of data²⁾ from various nuclides has been accomplished in the same way as Hughes and Harvey³⁾ did. The resulting curve is shown in Fig. 1.

It is rather easy to understand the exponential D distribution just shown, if it is assumed that *a priori random distribution* of levels is valid. Under the assumption just described, the problem of the size distribution is equivalent to that of distribution of the nearest neighbor⁵⁾ in a random

distribution of nuclear levels.

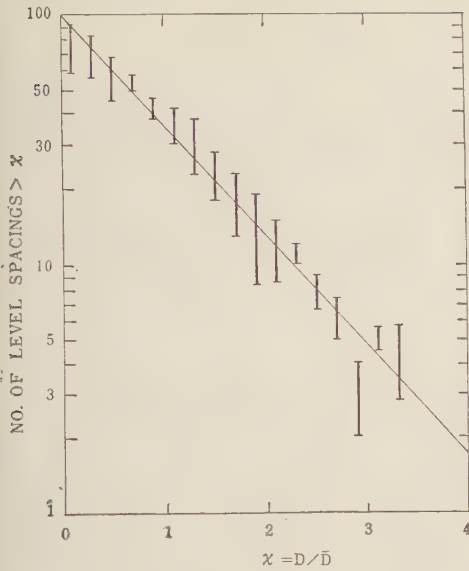


Fig. 1 The distribution of 268 level spacings relative to the average value for each nuclide. The number of level spacings exceeding a particular value is plotted against this value. The distributions for different nuclides are normalized so they can be plotted together.

Let $\omega(D)dD$ denote the probability that the nearest neighbor to a level occurs between D and $D+dD$. This probability must be clearly equal to the probability that no levels exist interior D times the probability that a level does exist in the range between D and $D+dD$. Then the function $\omega(D)$ must satisfy the relation

$$\omega(D) = [1 - \int_0^D \omega(D) dD] / \bar{D}, \quad (1)$$

where \bar{D} denotes the averaged level spacing, which is given by E/N : N is the number of levels and E is the energy interval containing N levels. From (1) we derive

$$\omega(D) = (1/\bar{D}) \exp[-D/\bar{D}]. \quad (2)$$

It should not be thought that the size distribution presented here is only restricted to the energy region just above the neutron

binding. The parameter which characterizes the size distribution is clearly the averaged level spacing, the inverse of which is just predicted by Weisskopf's formula

$$1/\bar{D} = c \exp(aE)^{1/2}. \quad (3)$$

Needless to say, this formula depends on the atomic weight through parameters a and c , and excitation energy. Thus, it is the second aim to show that the size distribution can be related intimately to atomic weight and excitation energy. Substituting (3) in (2), we obtain the dependence of the fluctuation of level spacing from the averaged one on atomic weight and excitation energy:

$$\omega(D) = c \exp(aE)^{1/2} \cdot \exp[-cD \exp(aE)^{1/2}]. \quad (4)$$

This relation may be useful, as it is possible to account for the fluctuation of level spacing in various excitation energy. For example, it results that at lower excitation the fluctuation is considerably large and at higher the smaller in accord with the variation of \bar{D} . Further details are under investigation.

One of the authors (S.O.) is indebted to the Yukawa-Yomiuri fellowship for the financial aid. The other author (F.F.) wishes to express his gratitude to Prof. M. Kobayasi, who afforded him the opportunity of staying at the Kyoto University, and to the Hyogo University of Agriculture for awarding the grant which made his stay possible.

- 1) J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (1952).
- 2) Harvey, Hughes, Carter and Pilcher, *Phys. Rev.* **99** (1955), 10.
D. J. Hughes and J. A. Harvey, *Neutron Cross Section* (1955) for In^{115} , Cs^{135} , Eu^{151} , Tb^{158} , Ho^{165} , Tm^{169} , Lu^{175} , Hf^{179} , Ta^{232} , U^{235} , and U^{238} .

- 3) D. J. Hughes and J. A. Harvey, Phys. Rev 99 (1955), 1032.
- 4) See, for example, Fig. 5 of reference 2).
- 5) See, for example, S. Chandrasekhar, Rev. Mod. Phys. 15 (1948), 1.

Note on the Decay of Ξ Particle

Shigeaki Hatano

Department of Physics, Kyoto University,
Kyoto

July 9, 1956

Recently Sakata¹⁾ proposed a compound hypothesis for the heavy unstable particles. According to this hypothesis, nucleons (N) and Λ particles and their antiparticles (\bar{N} and $\bar{\Lambda}$) are considered to be fundamental and other particles are composed of them in such a way as $\pi(N\bar{N})$, $K(N\bar{\Lambda})$, $\Sigma(\Lambda\bar{N})$ and $\Xi(\Lambda\bar{\Lambda})$ (each particle is composed of the particles in parentheses). We assume that a particle and an antiparticle have a tendency to be in a couple, by analogy with the deuteron model of atomic nucleus. Thus, in Sakata's model, we regard Ξ as a bound system of Λ and \bar{K} (anti- K particle) and Σ as Λ and π .

To present this idea in a mathematical form, we use a formalism of meson pair theory²⁾. Assuming that Λ is a particle of spin 1/2 and isospin 0 and \bar{K} is of spin 0 and isospin 1/2, a possible form of the strong, charge independent interaction between them is

$$H_{\Lambda\bar{K}} = -f_1 \int d\mathbf{x} \bar{\psi}_\Lambda(\mathbf{x}) \psi_\Lambda(\mathbf{x}) \phi^+(\mathbf{x}) \phi(\mathbf{x}), \quad (1)$$

where

$$\phi(\mathbf{x}) = \begin{pmatrix} \varphi_{\bar{K}^0} \\ \varphi_{\bar{K}^-} \end{pmatrix}.$$

Here we take the interaction attractive so that the bound state comes from this. If we follow the usual procedure in meson pair theory²⁾ and make the static approximation for a case of one Λ particle resting at the origin, (1) is written in momentum space as

$$H_{\Lambda\bar{K}} = - \frac{f_1}{(2\pi)^3} \int d\mathbf{k} d\mathbf{k}' v^*(\mathbf{k}) v(\mathbf{k}') \cdot \phi^+(\mathbf{k}) \phi(\mathbf{k}'),$$

where $v(\mathbf{k})$ is a cut-off factor. The equation of motion for $\phi(\mathbf{k})$ is

$$(E^2 - \kappa^2 - k^2) \phi(\mathbf{k}) + \frac{f_1}{(2\pi)^3} v^*(\mathbf{k}) \cdot \int d\mathbf{k}' v(\mathbf{k}') \phi(\mathbf{k}') = 0, \quad (2)$$

where κ is the mass of \bar{K} particle. The total energy E is determined by the secular equation

$$1 + \frac{f_1}{(2\pi)^3} \int d\mathbf{k} \frac{|v(\mathbf{k})|^2}{E^2 - \omega_k^2} = 0. \quad (3)$$

For a bound state, the total energy E should be

$$\kappa^2 - E^2 > 0.$$

The Ξ particle is represented as a bound state in our model when E is taken as

$$E = m_\Xi - m_\Lambda,$$

where m_Λ and m_Ξ are the masses of Λ and Ξ particles respectively. Using the cut-off factor

$$|v(\mathbf{k})|^2 = A^2 / (A^2 + k^2)$$

and taking a cut-off momentum $A = m_\Xi$, we obtain

$$f_1/4\pi = \frac{1}{A} \left(1 + \frac{W}{A} \right) = 1.3 \times \frac{1}{m_{\Xi}} \quad (4)$$

where $W = \sqrt{\kappa^2 - E^2}$.

The solution of Eq. (2) which belongs to the above eigenvalue E is

$$\varphi_0(\mathbf{k}) = \frac{1}{2\pi} \sqrt{\frac{2W}{E}} \left(1 + \frac{W}{A} \right) \frac{v^*(\mathbf{k})}{W^2 + k^2}. \quad (5)$$

In our model this wave function describes the state of Ξ particle. Of course in addition to these bound states ($\Lambda^0 \bar{K}^0$) and ($\Lambda^0 K^-$), other bound states ($\Lambda^0 K^+$) and ($\Lambda^0 \bar{K}^0$) arise in degenerating to the above states. But these redundant states will decay rapidly into ordinary particles, because they have strangeness 0.

In the case of Σ particle we take the interaction between Λ and π as

$$H_{\Lambda\pi} = -f_2 \int d\mathbf{x} \bar{\psi}_{\Lambda}(\mathbf{x}) \psi'_{\Lambda}(\mathbf{x}) \varphi_{\alpha}(\mathbf{x}) \varphi_{\alpha}(\mathbf{x}). \quad (6)$$

Calculating in a similar manner as before and putting $E = m_{\Sigma} - m_{\Lambda}$ and $A = m_{\Sigma}$, we obtain

$$f_2/4\pi = 1.1 \times (1/m_{\Sigma}). \quad (7)$$

Now we consider the decay process $\Xi^- \rightarrow \Lambda^0 + \pi^-$. In our model this process corresponds to a diagram in Fig. 1 (a) or 1 (b) according as we identify K^- particle with θ^- or τ^- meson.

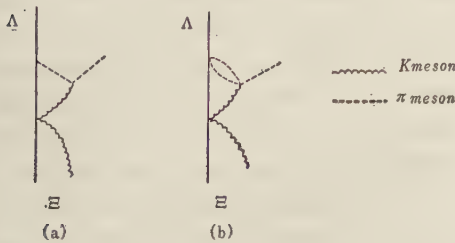


Fig. 1. Decay process of Ξ particle. (a) is through a decay of θ meson, and (b) is of τ meson.

It is found in Fig. 1 (a) that Λ can not absorb a single meson, because Λ has isospin 0 and π has isospin 1. Therefore the only possible process is the one with a decay of τ meson shown in Fig. 1 (b). We take the interaction which describes the decay of τ^- meson into 3π mesons as

$$g \int d\mathbf{x} \phi_{\tau^-}(\mathbf{x}) \varphi_{\alpha}(\mathbf{x}) \varphi_{\alpha}(\mathbf{x}) \varphi_{\pi^-}^*(\mathbf{x}) + \text{Complex Conjugate}. \quad (8)$$

Using the interactions (6) and (8) and the wave function (5), we calculate the transition probability of the decay process of Ξ particle shown in Fig. 1 (b) in the lowest order perturbation approximation. Then the ratio of the lifetime of τ^- meson to that of Ξ particle is estimated. If the value of f_2 in (7) is used, the result is

$$\tau_{\tau^-}/\tau_{\Xi} \sim 1 \times 10^3.$$

This is very interesting because the experimental data for the ratio show the same order of magnitude.

Although we neglect the recoil and the pair formation of Λ particles, the fact that the lifetime of Ξ^- particle can be explained qualitatively will indicate hopefulness of our model. In conclusion it is remarked that in our model Ξ and Σ have an opposite parity to that of Λ .

The author would like to express his sincere thanks to Professors M. Kobayasi, Y. Fujimoto and S. Takagi for their continual encouragement and stimulating discussions.

- 1) S. Sakata, Lecture at the Annual Meeting of the Physical Society of Japan, October, 1955.
- 2) E. P. Wigner, C. L. Critchfield and E. Teller, Phys. Rev. 56 (1939), 530.
G. Wentzel, Helv. Phys. Acta 15 (1942), 111.

On the Energy Dependence of the Cross Section for the Production of the Penetrating Shower Underground

S. Higashi, M. Oda, T. Oshio,
H. Shibata, K. Watanabe and
Y. Watase

*Institute of Polytechnics, Osaka City University,
Osaka*

June 18, 1956

It seems to be important to know the cross section for nuclear interactions produced by mu-mesons as a function of the incident energy in order to understand these phenomena.

There have so far been various experiments on the nuclear interactions underground (U.P.S.)¹⁾⁻⁵⁾ and their cross sections have been estimated for various depths of observation and energies transferred to the interactions. It was reasonable to assume that the major part of the incident particle of U.P.S. is the mu-meson.^{1), 3), 6)}

The cross section as a function of the depth of observation is to be related to the cross section as a function of the incident energy, E , of mu-meson by the following equation ;

$$\sigma_{\mu}(x, \varepsilon) = \int_{\varepsilon}^{\infty} \sigma_{\mu}(E, \varepsilon) |\partial I(x, E) / \partial E| \cdot dE \quad (1)$$

where ε : threshold energy of this event,
 $I(x, E)$: intensity of the hard component which has the energies higher than E at the depth x .

The cross section of this form is directly introduced from the experiment with the following expression ;

$$\sigma_{\mu}(x, \varepsilon) = n(x, \varepsilon) / I(x, 0) \cdot N \cdot X \quad (2)$$

where $n(x, \varepsilon)$: frequency of the nuclear interactions at the depth x , involving the energies greater than ε ,
 N : Avogadro's number,
 X : thickness of the shower producer in g/cm^2 .

The experimental cross sections obtained by various investigators are listed in Table I. $I(x, E)$ is obtained from the intensity-depth relation of the hard component underground and a presumed energy-range relation of the mu-meson.

George and Evans¹⁾ proposed an interpretation of the nuclear interaction underground as a process of photo-nuclear interaction by virtual photons associated with the mu-meson in flight and presented the following expression :

$$\sigma_{\mu}(E, \varepsilon) = (\alpha \sigma_{\gamma} / \pi) [\ln(E/\varepsilon)]^2 \quad (3)$$

where α : fine structure constant,

σ_{γ} : the cross section for photo-nuclear interaction of the virtual photon, and was presumed as $10^{-28} cm^2/nucleon$.

Although they were successful in interpreting their own experimental conclusions, the discrepancy between the experimental cross sections and the calculated values on the basis of their assumption seems to be appreciable beyond the statistical errors as listed in Table I, and therefore, it is hard to say that the above interpretation affords great satisfaction to the total cross sections over the wide region of the threshold energies and depths.

It may be worth-while to attempt to make a semi-empirical expression of $\sigma(E, \varepsilon)$ which is consistent with experimental results in the form (2). It seems to be necessary to introduce at least two parameters into

Table 1

| Observer | Depth* (m.w.e.) | Cross Section (cm ² /nucleon) | E_0^{**} | Cross Section Estimated by the Weizsäcker-Williams Method*** (cm ² /nucleon) |
|----------------|--------------------|---|----------------|---|
| George & Evans | 30 | $0.82 \pm 0.4 \cdot 10^{-30}$ | ≥ 200 MeV | 7×10^{-30} |
| | 50 | $2.6 \pm 0.8 \cdot 10^{-30}$ | " | |
| | 68 | $1.35 \pm 0.4 \cdot 10^{-30}$ | " | |
| | Mean | $1.54 \pm 0.3 \cdot 10^{-30}$ | " | |
| Cornell Group | 1600 | $4 \pm 2 \cdot 10^{-29}$ | ≥ 5 GeV | 6×10^{-30} |
| Osaka Group | 32 | $2.8 \pm 0.7 \cdot 10^{-31}$ | ≥ 7 GeV | 2.8×10^{-31} |
| | 200 | $2.9 \pm 1.1 \cdot 10^{-30}$ | " | 1.7×10^{-30} |

Table I: Experimental cross section for penetrating shower production by various authors.

*: from the top of the atmosphere

**: the energy involved in the shower estimated on the assumption that the nature of the shower underground is essentially similar to that at mountain altitude

***: on the assumption of constant cross section, 10^{-28} cm²/nucleon, for photo-nuclear interactions

the expression (3) to cover the above experiments, and the following expression was adopted:

$$\sigma_{\mu}(E, E_{\min}) = (2\alpha/\pi) \int_{E_{\min}}^E \sigma_{\gamma} \ln(E/E') dE'/E', \quad (4)$$

$$\sigma_{\gamma} = k_1 \beta^2 (E')^3,$$

$$E_{\min} = \eta E_0$$

where E : threshold energy of the showers estimated from the experimental results at the mountain altitude, assuming that U.P.S.'s have the same mechanism as that of showers observed at mountain altitude.

β and η may correspond to the following physical meanings: β representing the degree of the energy dependence of σ_{γ} , and $1/\eta$ being roughly referred to the inelasticity. With (1) and Table I one reaches the result that $\beta \simeq 0.6$ and $\eta \simeq 2$. These lead to a semi-empirical expression of the cross section as follows;

$$\sigma_{\mu}(E, E_{\min}) = (2ak/\pi) [E^{0.6} - E_{\min}^{0.6} \cdot (0.6 \ln(E/E_{\min}) + 1)] \quad (5)$$

where E_{\min} is about twice as large as the threshold energy of the showers of the same sizes at mountain altitude. The cross section of U.P.S., as a result, has an energy dependence which is slightly stronger than that led by Weizsäcker-Williams method⁷⁾ which was adopted in the interpretation by George and Evans.

The authors are grateful to the Institute for Food Chemistry for the financial aid.

- 1) E. P. George and J. Evans, Proc. Phys. Soc. **A63** (1950), 1248; **A64** (1951), 193; **A68** (1955), 829.
- 2) E. Amaldi, C. Castagoli, A. Gigli and S. Sciuti, Nuovo Cimento **9** (1952), 453; **9** (1952), 969. P. E. Argan, A. Gigli and S. Sciuti, Nuovo Cimento **11** (1954), 530.
- 3) P. H. Barret, M. L. Bollinger, G. Cocconi, Y. Eisenberg and K. Greisen, Rev. Mod. Phys. **24** (1952), 133.
- 4) S. Higashi, I. Higashino, M. Oda, T. Oshio, H. Shibata, K. Watanabe and Y. Watase, J. Phys. Soc. Japan, under publication.
- 5) G. Cocconi and V. Cocconi-Tongiorgi, Phys. Rev. **84** (1951), 29.
- 6) S. Higashi, M. Oda, T. Oshio, H. Shibata and K. Watanabe, J. Phys. Soc. Japan **9** (1954), 655.
- 7) E. J. Williams, Kgl. Danske Vidensk. Selsk. **13** (1935), No. 4.

On the Nuclear Interaction of μ -Meson below Ground

Takashi Kitamura and Minoru Oda

*Institute of Polytechnics,
Osaka City University, Osaka*

June 18, 1956

There have been several investigations^{1), 2), 3)} on the nuclear interactions underground. George and Evans¹⁾ and other investigators⁴⁾ had shown that they could be interpreted as photo-nuclear interactions by virtual photons associated with μ -mesons in flight. It has been shown by Higashi et al.⁵⁾ that, if one would cover experimental conclusions at various depths underground by this interpretation, one reaches a cross section for the photo-nuclear interaction as a sort of increasing function of the energy of photon.

So far, the role of secondary particles of the interactions has been neglected and, hence, the interactions have been considered to be wholly produced by μ -mesons. The present authors would like to show briefly that the proportion of the nuclear interaction underground produced by secondary π -meson component is possibly considerable so long as one takes the above interpretation with the assumption that the nature of the nuclear interaction of μ -meson is essentially similar to that of the nuclear interaction at high altitude, that is, the nuclear interaction produced by the nucleonic component.

The number of π -mesons at an energy interval $(E_\pi, E_\pi + dE_\pi)$ at the depth, x , are expressed as a function $f(E_\pi, x)$ which satisfies the following equation;

$$\frac{\partial f(E_\pi, x)}{\partial x} = \eta \frac{\partial f(E_\pi, x)}{\partial E_\pi} - \frac{1}{\lambda} f(E_\pi, x) - \frac{\mu}{c E_\pi \tau_\pi} f(E_\pi, x) + G(E_\pi, x), \quad (1)$$

η : ionization loss; 2.2 Mev/g/cm² in rock,

λ : mean free path for absorption;

200 g/cm² in rock,

ρ : density of rock,

τ : mean lifetime of π -meson,

μ : rest energy of π -meson,

$G(E_\pi, x) dE_\pi dx$: the number of π -mesons produced at the depth between x and $x + dx$ whose energy is in the interval $E_\pi, E_\pi + dE_\pi$.

Solving the equation with a condition that $f(E_\pi, x) = 0$ at $x = 0$, it turns out⁶⁾ to be that:

$$f(E_\pi, t) = \int_0^t G(E_\pi + \eta(t-x)) \exp\left(-\frac{t-x}{\lambda}\right) \left(1 + \frac{\eta(t-x)}{E_\pi}\right)^{-\mu/\eta c \rho \tau} dx. \quad (2)$$

If one takes the interpretation following George and Evans that a μ -meson is equivalent to a group of virtual photon, $G(E_\pi, x)$ is to be expressed as follows:

$$G(E_\pi, x) dE_\pi = dE_{ph} \int_E^\infty I(x, E_\mu) P(E_\mu, E_{ph}) \cdot \sigma_{ph-N}(E_{ph}) n(E_{ph}) N dE_\mu, \quad (3)$$

where

$I(x, E_\mu)$: the energy spectrum of μ -mesons at the depth, x ,

$P(E_\mu, E_{ph})$: the energy spectrum of virtual photons, E_{ph} , associated with a μ -meson of energy, E_μ . The expression of Williams and Weizsäcker is to be used here following George and Evans,

$\sigma_{ph-N}(E_{ph})$: the cross section for the photo-nuclear interaction by a photon of

energy E_{ph} ,

$n(E_{ph})$: average multiplicity of charged π -mesons produced by a photonuclear interaction as a function of the energy of photon, E_{ph} ,

N : Avogadro number,

ε : threshold energy for a penetrating shower produced by μ -meson.

E_π and E_{ph} can be related, on the average, with each other, provided that $n(E_{ph})$ is known. The relation, $n(E_{ph})$, between the multiplicity of produced π -mesons and the energy of photon, is presumably assumed to be similar to that between the multiplicity and the energy of nucleonic component in the case of nuclear interaction at high altitude. It is postulated, following George and Evans, that $\sigma_{ph-N}(E_{ph}) = \text{const} \approx 10^{-28} \text{ cm}^2$.

The results of calculation at 32 m.w.e., 200 m.w.e. and 1600 m.w.e. below the top of the atmosphere are represented in the figure. The cross is a value estimated from the ratio of the number of slow π -mesons and μ -meson coming to rest in the emulsions exposed underground.⁷⁾ With these energy spectra of π -mesons underground and presumed cross section, 350 mb, for the nuclear interaction of π -mesons, one can estimate the relative frequency of occurrence of nuclear interactions underground for various depths and threshold energies of the interactions as shown in the table. Although it should be emphasized that the arguments are rather indefinite because of lack of our knowledge on the nuclear interaction underground, the role of secondary π -mesons in the nuclear interaction underground seems to be considerable particularly at larger depths unless a radically different mechanism on the nuclear interaction under-

ground is to be adopted.

We wish to express our thanks to Prof. Y. Watase for suggesting this work as well as for his continuous interest in the subject. Thanks are also due to the μ -meson Group^{*)} in this laboratory for their discussion.

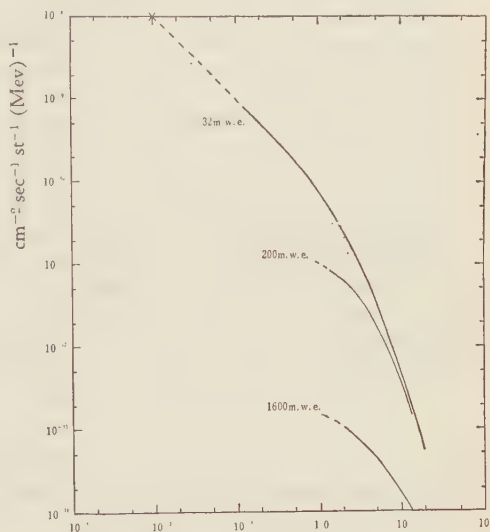


Fig. 1. E_π in Bev

Energy spectrum of secondary π -mesons.

Table 1.

| threshold energy | 1 Bev | 3 Bev |
|---------------------|-------|-------|
| depth | | |
| 32 m.w.e. | 0.7 | 0.5 |
| | 1.0 | 1.0 |
| 200 m.w.e. | 1.75 | 1.1 |
| | 1.0 | 1.0 |
| 1600 m.w.e. | 2.45 | 1.75 |
| | 1.0 | 1.0 |

Relative frequency of occurrence of the nuclear interactions underground produced by π -mesons to that produced by μ -mesons.

*) S. Higashi, T. Oshio, H. Shibata, K. Watanabe, S. Fukui and Y. Murata

- 1) E. P. George and J. Evans, *Proc. Phys. Soc.* **A63** (1950), 1248; *ibid.* **A64** (1951), 193.
- 2) E. Amaldi, C. Castagnoli, A. Gigli and S. Scuti, *Nuovo Cim.* **9** (1952), 453; 969.
A. Lovati, A. Mura, C. Succi and G. Tagliaferri, *Nuovo Cim.* **10** (1953), 103; 1201.
P. E. Argan, A. Gigli and S. Scuti, *Nuovo Cim.* **11** (1954), 530.
H. J. J. Braddick and B. Leontic, *Phil. Mag.* **45** (1954), 1287.
- 3) S. Higashi, I. Higashino, M. Oda, T. Oshio, H. Shibata, K. Watanabe and Y. Watase, to be published.
- 4) P. H. Barret, M. L. Bollinger, G. Cocconi, Y. Eisenberg and K. Greisen, *Rev. Mod. Phys.* **24** (1952), 133.
- 5) S. Higashi, M. Oda, T. Oshio, H. Shibata, K. Watanabe and Y. Watase, the preceding letter.
- 6) C. Franzinetti, *Nuovo Cim.* **73** (1950), 84.
- 7) E. P. George and J. Evans, *Proc. Phys. Soc.* **A68** (1955), 829.

Thermodynamic Properties of He^3

V. S. Nanda and S. K. Trikha

*Department of Physics, Delhi University,
Delhi, India*

July 5, 1956

It is natural to attempt an explanation of the properties of He^3 on the basis of models similar to those which have been tried in the case of He^4 , for both are essentially quantum liquids. In the beginning some authors^{1),2),3)} applied the perfect F.D. gas model to liquid He^3 . The measurement of its specific heat later on, however, gave values very much different from the theoretical results. Further, the nuclear spin magnetic susceptibility, as measured by Fairbank et al.⁴⁾, was found

to be in complete disagreement with theory. This led Price⁵⁾, Rice⁶⁾ and Temperley⁷⁾ to give up the above approach altogether and use the 'cell model' of the liquid.

The object of the present note is to show that the possibilities of a F.D. model for He^3 have not been altogether exhausted. In fact, a F.D. gas model, modified on the lines of a corresponding model for He^4 , leads to a fairly satisfactory explanation of the thermodynamic behaviour of He^3 . An inevitable consequence of a model of the above type is the existence of partial nuclear spin alignment in the liquid at all temperatures.

Following London⁸⁾ in the case of liquid He^4 , we describe liquid He^3 by picturing it as a metal in which the ions and the electrons are replaced by particles of the same kind viz. He^3 atoms. A fraction $f(=1/3.8)$ of the total number of atoms corresponds to the free electrons in the metal and behaves as an ideal F.D. gas in a potential well of depth Z , while the rest of the atoms constitute a quasi-crystalline lattice. In our computations of the specific heat and entropy, we have taken the F.D. degeneracy temperature to be 0.54°K and a Debye temperature 8.5°K . In Fig. 1 and Fig. 2 we show a comparison between the theoretical and experimental specific heat and entropy values^{9),10)}. The agreement below 2°K is fairly satisfactory. The discrepancy above this temperature can be attributed to the lowering of the Debye temperature (see in this connection Mikura¹¹⁾).

It is significant to note that in our computations we have not taken into account the contribution to the total entropy of the liquid arising out of the possible random alignment of the nuclear spins of

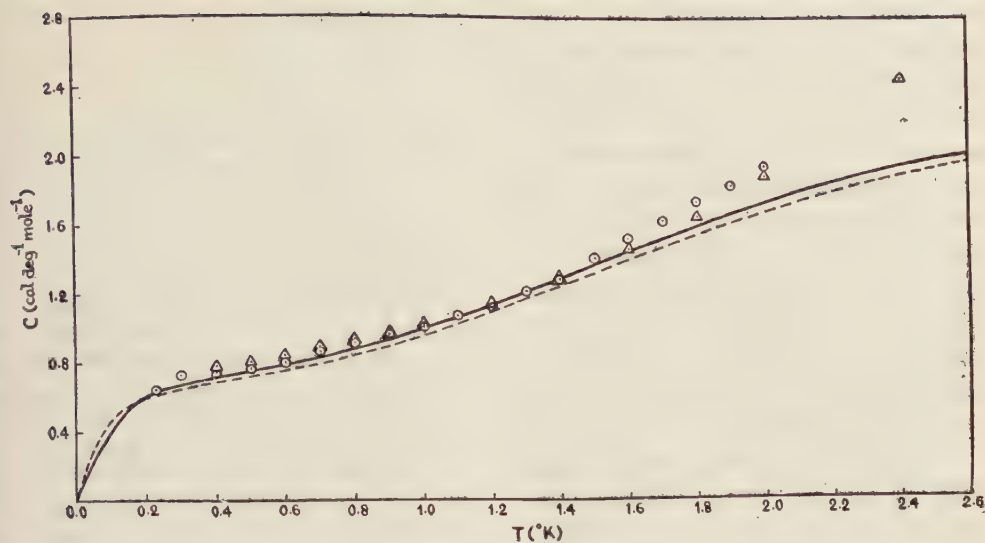


Fig. 1. The specific heat, C cal. mole⁻¹ deg⁻¹, of liquid He³ as a function of absolute temperature, T (°K).

—: Theoretically calculated values with $T^* = 0.54^\circ\text{K}$, $f = 1/3.8$.

- - -: Theoretically calculated values with $T^* = 0.45^\circ\text{K}$, $f = 1/4.0$.

○: Smoothed experimental results of Abraham et al.⁹⁾

△: Experimental results of Roberts and Sydorik.¹⁰⁾

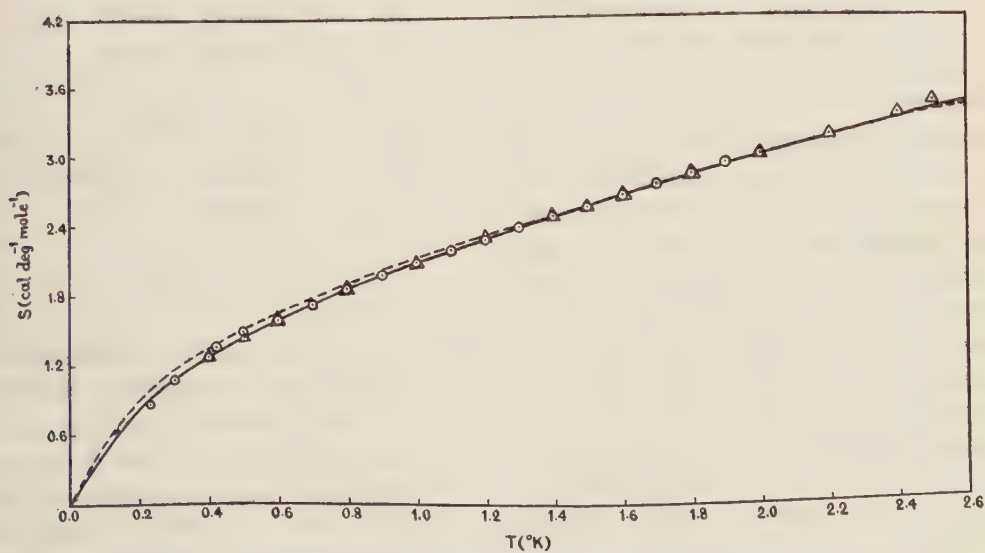


Fig. 2. The entropy, S cal. mole⁻¹ deg⁻¹, of liquid He³ as a function of absolute temperature, T (°K).

—: Theoretically calculated values with $T^* = 0.54^\circ\text{K}$, $f = 1/3.8$.

- - -: Theoretically calculated values with $T^* = 0.45^\circ\text{K}$, $f = 1/4.0$.

○: Smoothed experimental results of Abraham et al.⁹⁾

△: Experimental results of Roberts and Sydorik.¹⁰⁾

the atoms which constitute the quasi-crystalline lattice. Thus, according to the present model the nuclear spins of these atoms are aligned for all liquid temperatures and therefore make no contribution to the nuclear paramagnetic susceptibility. The measured susceptibility is, therefore, due only to the fraction which behaves as an ideal gas with a degeneracy temperature of 0.54°K . This is very close to the degeneracy temperature of 0.45°K required by Fairbank et al. to fit their experimental results with the theoretical F.D. curve.

In the end it appears worthwhile to point out how the various arbitrary parameters θ , f and T^* occurring in the theory have been fixed. The choice of $\theta (=8.5^\circ\text{K})$ is governed by the shape of the specific heat curve above 1°K and its value lies between the range 9.25 and 7.52°K arrived at by Mikura¹¹⁾ by applying the law of corresponding states. The parameter $f(=1/3.8)$ is fixed from the shape of the specific heat curve below 1°K while $T^*(=0.54^\circ\text{K})$ is chosen so as to give the observed value of the entropy^{9),10)} at 1.5°K . Fairbank et al., however, found that their susceptibility data could be fitted in with the susceptibility of an ideal F.D. gas with $T^*=0.45^\circ\text{K}$. We have, therefore, shown in Fig. 1 and Fig. 2 the entropy and specific heat computations with this value of T^* (dotted curves). Here the original value of θ is retained while f is taken as equal to $1/4.0$. The agreement in this case is not so good as compared with our first choice, but is still in reasonable accord with experiment.

A detailed discussion of the consequences of the present model is deferred to a later publication.

Our grateful thanks are due to Prof.

D. S. Kothari for his interest in this investigation.

- 1) K. S. Singwi, Phys. Rev. **87** (1952), 540.
- 2) I. Pomeranchuk, J. Exp. Theor. Phys. (USSR) **20** (1950), 919.
- 3) L. Goldstein, (unpublished).
- 4) W. M. Fairbank, W. B. Ard and G. K. Walters, Phys. Rev. **95** (1954), 566.
- 5) P. J. Price, Phys. Rev. **97** (1955), 259.
- 6) O. K. Rice, Phys. Rev. **97** (1955), 263.
- 7) H. N. V. Temperley, Phys. Rev. **97** (1955), 835; Proc. Phys. Soc. **68A** (1955), 1136.
- 8) F. London, J. Phys. Chem. **43** (1939), 49.
- 9) B. M. Abraham, D. W. Osborne and B. Weir, Phys. Rev. **98** (1955), 551.
- 10) T. R. Roberts and S. G. Sydorak, Phys. Rev. **98** (1955), 1672.
- 11) Z. Mikura, Prog. Theor. Phys. **13** (1955), 120.

On the Energy Matrix of the Tensor Forces

Tatuya Sasakawa

*Department of Physics, Kyoto University,
Kyoto*

July 2, 1956

In investigating nuclear configurations, there arises a need for evaluating the two particle matrix elements in $j-j$ or $L-S$ coupling scheme. In this note our attention will be directed to the matrix elements of the two particle tensor force operators between the states of two equivalent or inequivalent nucleons.

In the $L-S$ coupling scheme, it is convenient to express the interaction of the tensor force between nucleons 1 and 2,

$$S_{12}V_{12} = \left[\frac{(\mathbf{s}_1 \cdot \mathbf{r}_{12})(\mathbf{s}_2 \cdot \mathbf{r}_{12})}{r_{12}^2} - \frac{1}{3}(\mathbf{s}_1 \cdot \mathbf{s}_2) \right] \cdot V(\mathbf{r}_{12}), \quad (1)$$

as

$$[V_{12}(\mathbf{r}_{12})/r_{12}^2] \cdot (\mathbf{s}_1 \cdot \mathbf{r}_{12})(\mathbf{s}_2 \cdot \mathbf{r}_{12}) - \frac{1}{3}(\mathbf{s}_1 \cdot \mathbf{s}_2)V(\mathbf{r}_{12}), \quad (2)$$

as was done by Talmi¹⁾ and recently by Hope and Longdon²⁾.

But in the j - j coupling scheme it is not convenient to work with the double tensors resulting from the expression (2), while they are convenient in the L-S coupling scheme. Talmi¹⁾ first noted this, and he showed that the expression (1) for $S_{12}V_{12}$ can be expressed as a sum of scalar products of tensors of degree K , irreducible with respect to \mathbf{j}_i :

$$S_{12}V_{12} = \sum_{kk'K} A_{kk'K} (\mathbf{t}_1^{(1k)K}, \mathbf{t}_2^{(1k')K}). \quad (\text{Talmi}^1) \quad (28)) \quad (3)$$

Notations used here are the same as Talmi's.¹⁾ Using this form, he evaluated the matrix elements of $(d_{5/2})^2$ and $(d_{5/2})^3$ configuration in terms of the Talmi integral.

However, he did not give the explicit expression of $A_{kk'K}$ and did not consider the K dependence of $A_{kk'K}$ in (3). To get the accurate evaluation for matrix element, $A_{kk'K}$ must be given in the explicit form. For this purpose we reduced $S_{12}V_{12}$ to the form

$$S_{12}V_{12} = \sum_{kk'K} A_{kk'K}^K (\mathbf{t}_1^{(k)K}, \mathbf{t}_2^{(k')K}), \quad (4)$$

where

$$A_{kk'K}^K = \sum_t 5 \sqrt{15/2\pi} (-)^{K+k-k'} \cdot \sqrt{(2k+1)(2k'+1)(2t00|k0)(2t00|k0)} \cdot W(k1k'1:K2)W(k2k'2:t2) \cdot v_t(\mathbf{r}_1\mathbf{r}_2), \quad (5)$$

$v_t(\mathbf{r}_1\mathbf{r}_2)$ being the radial part of V_{12} , when it is expressed as

$$V_{12} = \sum_l v_l(r_1r_2) (\mathbf{c}_1^{(l)} \cdot \mathbf{c}_2^{(l)}), \quad (6)$$

$$c_{i,\rho}^{(l)} = [4\pi/(2l+1)]^{1/2} Y_\rho^{(l)}(\theta_i, \varphi_i). \quad (7)$$

The matrix element of $(\mathbf{t}_1^{(1k)K}, \mathbf{t}_2^{(1k')K})v_t$ is generally expressed using the U -coefficient⁵⁾ as,

$$\begin{aligned} & \langle \varepsilon j_n j X || (\mathbf{t}_1^{(1k)K}, \mathbf{t}_2^{(1k')K}) v_t || \varepsilon' j_n' j' X \rangle \\ &= (-)^{j_n+j'-X} W(j_n j_n' j' X) \\ & \cdot \sqrt{(2j_n+1)(2j_n'+1)(2j+1)(2j'+1)} \\ & \cdot (2K+1) \\ & \cdot \langle s || \sigma^{(1)} || s \rangle \langle s || \sigma^{(1)} || s \rangle \langle l_n || c^{(k)} || l_n' \rangle \langle l || c^{(k')} || l' \rangle \\ & \cdot U \begin{pmatrix} s & l_n & j_n \\ s & l_n' & j_n' \\ 1 & k & K \end{pmatrix} U \begin{pmatrix} s & l & j \\ s & l' & j' \\ 1 & k' & K \end{pmatrix} \iint R_{l_n} R_{l_n'} R_l R_{l'} \\ & \cdot v_t(12) r_1^2 dr_1 r_2^2 dr_2. \end{aligned} \quad (8)$$

Using (8) we can calculate the matrix elements between various configurations with the technique due to Racah⁴⁾.

Particularly in the case or $l_n=l_n'$ and $j_n=j_n'$, a U -coefficient in (8) vanishes unless $1+k+K=\text{even}$. Hence K must be odd because k is always even owing to the parity conservation.

A direct consequence of the above argument is as follows. In a matrix element

$$\langle j_n^p(I)j : \varepsilon JM | V_{12} S_{12} | j_n^p(I')j : \varepsilon' JM \rangle \quad (9)$$

it can be shown $I'=K$ if $I=0$. Then the matrix element (9) vanishes whenever $I=0$ and $I'=\text{even}$. This conclusion includes the vanishing of tensor force interactions between a closed shell and an external nucleon, also between a closed shell and an almost closed shell, as special cases.

- 1) I. Talmi, Phys. Rev. **89** (1953), 1065.
- 2) J. Hope and L. W. Longdon, Phys. Rev. **101** (1956), 710.
- 3) Arima, Horie and Tanabe, Prog. Theor. Phys. **11** (1954), 143.
- 4) G. Racah, Phys. Rev. **62** (1942), 438, 5.

JOURNAL OF THE PHYSICAL SOCIETY OF JAPAN

Volume 11, Number 7, July 1956

CONTENTS

| | Page |
|---|------|
| Hisao YAGI: Transformations of Various Centres among Each Other in Coloured Alkali-halide Crystals | 723 |
| Toshio MISAWA: A Note on the Extended Theory of the Junction Transistor..... | 728 |
| Yasuyoshi WATANABE: On the Electron Diffraction Patterns from the Oriented Silver Films | 740 |
| Norio KATO: Anomalous Enhancement of X-ray Reflection Intensity at the Boundary of Ground and Etched Regions on Crystal Surface | 748 |
| Hachiro KOMATSU: Proton Magnetic Resonance Absorption Spectra of Long Chain Normal Primary Alcohol of Probable Carbon Number 32 | 755 |
| Tutô NAKAMURA and Norikiyo URYÛ: On the Interionic Contributions to the Specific Heat of Cobalt Ammonium Tutton Salt $\text{Co}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ at Very Low Temperatures | 760 |
| Norikiyo URYÛ: Note on the Magnetic Anisotropy of the Cobalt Ammonium Tutton Salt $\text{Co}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ | 770 |
| Kiyoshi MURAKAWA: Hyperfine Structure of the Spectrum of Cu I and Test of the Hyperfine Structure Formulas | 774 |
| Kiyoshi MURAKAWA: Isotope Shift in the Spectra of Ne II and W I..... | 778 |
| Tetsuji NISHIKAWA: Fine Structure of $J=1 \leftarrow 0$ Transition due to Internal Rotation in Methyl Alcohol | 781 |
| Hajime NARUMI: On the Eigenvalue Problem in Terms of a Complete Set of the Casimir Operators | 786 |
| Hidehori HASIMOTO: A Sphere Theorem on the Stokes Equation for Axisymmetric Viscous Flow | 793 |
| Yoshio SHIBAKA: On the Transverse Vibration of an Elliptic Plate with Clamped Edge | 797 |

SHORT NOTES

| | |
|--|-----|
| Shoichiro NOMURA: On a Change of Dielectric Property of $(\text{Ba}_{0.5}\text{Sr}_{0.5})\text{TiO}_3$ Ceramics due to the Thermal Treatment | 803 |
| Kiyoshi MURAKAWA: Nuclear Spins of Dy 163 and Dy 161 | 804 |
| Shoji KOJIMA, Kiyoe KATO and Shoichi ANNAKA: Radio Frequency Absorption in Barium Chlorate | 805 |

Errata

| | |
|--|-----|
| Yûkichi NOMURA and Shigetoshi KATSURA: Diffraction of Electromagnetic Waves by Circular Plate and Hole | 806 |
| Alan K. HEAD: The Effect of Frequency on the Fatigue of Metals | 806 |

The Range of the Interaction between Different Fermions

Michiji KONUMA

Department of Physics, University of Tokyo, Tokyo

(Received December 9, 1955)

We investigate the range of a virtual boson field coming from the transition between two fermions whose masses are different from each other. The spread of a boson emitted by a heavier fermion is longer than that of the one emitted by a lighter fermion. The range of the exchange force potential between two fermions of different masses is longer than the Compton wave length of the boson. The effects of this range shift on the interaction between hyperons and heavy mesons are considered.

§ 1. Introduction

It is well known that the ranges of both nuclear force and π meson cloud around the nucleon are given by the Compton wave length of the π -meson. It seems worthwhile to investigate the range of a virtual boson field arising from the transition between two fermions whose masses are different from each other,* since it will contribute to the construction of a quantitative theory on the interaction between hyperons and heavy mesons.

In § 2 and § 3, it is discussed that the range of the exchange force potential between two fermions of different masses is longer than the Compton wave length of the boson. In § 4, considering the effect on the heavy particles, we find that the effect of the range shift in comparison with the recoil effect is small in the case of $NA\theta$ interaction, etc., but is large in the case of $A\Sigma\pi$ interaction.

§ 2. A qualitative consideration

We start our discussions by a qualitative consideration of the interaction between two fermions F_1 , F_2 and a boson B , whose masses are denoted by m_1 , m_2 and μ respectively. In the case of

$$\mu > m_2 - m_1 \equiv \delta > 0, \quad (1)$$

a free fermion can not emit the real boson, but it can emit virtual bosons within the limit of the uncertainty principle.

We consider the process $F_1 \rightarrow F_2 + B$. Suppose that we observe the boson by setting up some device at a distance r from the fermion source. The time which the boson

* After writing this manuscript, the author was informed that Dr. A. Sugie had remarked the effect of the mass difference between charged and neutral bosons on the nuclear force (Prog. Theor. Phys. 11 (1954), 333).

takes to propagate this distance with the velocity v is given by

$$t = r/v \gtrsim r/c.$$

The energy shift ΔE of the system must be at least

$$\Delta E \gtrsim (\mu + m_2 - m_1)c^2 = (\mu + \delta)c^2.$$

Considering the uncertainty relation for energy and time and the consistence with the energy conservation law, the time Δt during which the boson exists in the system should be

$$\Delta t \lesssim \hbar/\Delta E \lesssim \hbar/(\mu + \delta)c^2.$$

Since the time of propagation must be shorter than Δt , we have

$$r < \hbar/(\mu + \delta)c. \quad (2)$$

Thus, the range of the boson field can not exceed $\hbar/(\mu + \delta)c$. In the same way the region of the boson emitted by the process $F_2 \rightarrow F_1 + B$ is limited by

$$r < \hbar/(\mu - \delta)c. \quad (2')$$

Therefore, the virtual boson field created by the transition between two fermions of different masses has two ranges, and the spread of the boson emitted by the heavier fermion is longer than that of the one emitted by the lighter fermion.

§ 3. The potential between different fermions

The above result is nothing but the upper limit of the two ranges. Hence, let us derive the potential between F_1 and F_2 . As an example, we take the interaction between two fermions—the nucleon (charge +, 0) and the Λ particle (charge 0)—and a scalar θ meson (charge +, 0) [each particle has its antiparticle]. The interaction Hamiltonian can be written as

$$H = g(\bar{\psi}_N \psi_N \phi_\theta + \text{herm. conj.}) \quad (3)$$

where

$$\bar{\psi}_N \phi_\theta = \bar{\psi}_p \phi_{\theta+} + \bar{\psi}_n \phi_{\theta 0}.$$

We can easily extend the calculation to the case of other types of the interaction, but the conclusion is expected to be unaffected. In the lowest order perturbation theory, the proton- Λ potential V is given, not including the recoil effects of fermions, by

$$\left. \begin{aligned} V &= V_a + V_b, \\ V_a &= -\frac{g^2}{(2\pi)^3} \int \frac{1}{\omega(-\omega - \delta)} \exp(ikr) dk, \\ V_b &= -\frac{g^2}{(2\pi)^3} \int \frac{1}{\omega(-\omega + \delta)} \exp(ikr) dk, \end{aligned} \right\} \quad (4)$$

where suffices a, b correspond to respective graphs in Figure 1 and hereafter we take

$\hbar=c=1$. Here, for convenience of the calculation, putting

$$\mu'^2 = \mu^2 - \partial^2, \quad (5)$$

$$\begin{aligned} I_1 &= -\frac{1}{(2\pi)^3} \int \frac{1}{\omega^2 - \partial^2} \exp(i\mathbf{k}\mathbf{r}) d\mathbf{k} \\ &= -\frac{1}{4\pi r} \exp(-\mu' r), \end{aligned} \quad (6)$$

$$I_2 = -\frac{\partial}{(2\pi)^3} \int \frac{1}{\omega(\omega^2 - \partial^2)} \exp(i\mathbf{k}\mathbf{r}) d\mathbf{k}, \quad (7)$$

the potentials have the forms

$$\left. \begin{aligned} V_a &= -(g^2/2)(I_1 - I_2), \\ V_b &= -(g^2/2)(I_1 + I_2). \end{aligned} \right\} \quad (4')$$

If we remind

$$-\frac{1}{(2\pi)^3} \int \frac{1}{\omega^3} \exp(i\mathbf{k}\mathbf{r}) d\mathbf{k} = (1/2\pi^2) K_0(\mu r)$$

with the Hankel function of imaginary argument $K_0(\mu r)$, it is found that

$$(\partial/2\pi^2) K_0(\mu' r) > I_2 > (\partial/2\pi^2) K_0(\mu r).$$

The $K_0(x)$ tends to infinity slower than $\exp(-x)/x$ for $x \rightarrow 0$, and for increasing x , it decreases monotonously and tends to zero as $\exp(-x)/x^{1/2}$ for $x \rightarrow \infty$. Thus I_2 has a larger spread than Yukawa potential I_1 , and for $\mu' r = 1$ the value of I_2 is 0.25 times as much as that of I_1 . Therefore, the range of $V_a(V_b)$ is of the value shorter (longer) than μ'^{-1} by the contribution from the second term in Equation (4'). (The range of V_a is still shorter than the Compton wave length μ^{-1} of the θ meson.) This conclusion is in agreement with the qualitative consideration in § 2.

As the static potential is given by

$$V = -\frac{g^2}{4\pi r} \exp(-\mu' r), \quad (8)$$

the range of the potential between two fermions of different masses is not the Compton wave length of the boson μ^{-1} , but

$$\mu'^{-1} = \{\mu^2 - (m_2 - m_1)^2\}^{-1/2}. \quad (9)$$

§ 4. The effects of the range shift on the heavy particles

Next, we consider qualitatively the effects of the range shift $\mu^{-1} \rightarrow \mu'^{-1}$ on the interactions of the heavy particles. According to the Nishijima²⁾ and Gell-Mann³⁾ model, the couplings in Table 1 are the strong, charge-independent interactions. (Moreover, if we take into account the pair interactions, $\Lambda\Lambda\pi\pi$ and $N\Xi\theta\theta$, etc., should be added.)

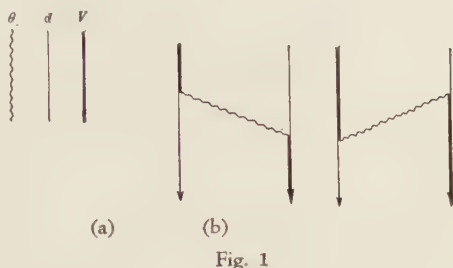


Table 1. Strong interactions.

| coupling | Compton wave length μ^{-1} | μ'^{-1}/μ^{-1} |
|--------------------|---------------------------------|----------------------|
| $NN\pi$ | $1.41 \times 10^{-13}\text{cm}$ | 1.00 |
| $\Lambda\Sigma\pi$ | | 1.19 |
| $\Sigma\Sigma\pi$ | | 1.00 |
| $\Xi\Xi\pi$ | | 1.00 |
| $N\Lambda\theta$ | $4.00 \times 10^{-14}\text{cm}$ | 1.07 |
| $N\Sigma\theta$ | | 1.08 |
| $\Lambda\Xi\theta$ | | 1.09 |
| $\Sigma\Xi\theta$ | | 1.04 |

In general, the recoil effect and the higher order corrections, etc., give various contributions to the interaction between two fermions. Hence, we are not necessarily allowed to consider only the range shift. Roughly speaking, for instance, in the case of $N.l^{\theta}$ interaction, while the inclusion of the recoil effect gives the contribution of the order $\mu/M \sim 1/2$ (M : the mean fermion mass), the effect of the range shift is of the order $\delta/\mu \sim 1/3$. However, in the case of $\Lambda\Sigma\pi$ interaction, it is expected that the effect of the range shift is not smaller than the others.

Recently, Nishijima^{2,11)} and Iwao⁶⁾ have discussed the $N.l$ bound system by using the Yukawa potential of the range μ_0^{-1} . In the case of the nuclear force, the potential of the meson theory is valid only in the outer region beyond the order of the half of π meson Compton wave length. About the inner region the validity of the potential concept is not yet clear. While the existence of the bound state depends on the volume of the inner region. Therefore, in the present case of the $N.l^{\theta}$ interaction it is questionable to connect the coupling constant of the Yukawa potential in the second order perturbation theory directly to the strength of the $N.l^{\theta}$ interaction. Hence we do not necessarily draw a definite conclusion, but if it is justified to use the Yukawa potential, we may say that, taking account of the range shift, the $N.l$ force becomes of the longer range than μ_0^{-1} , as a result the coupling between $N\Lambda\theta$ tends to be weaker.

To take into account the weak interactions relating to the decay mechanisms induces some new problems. Taking the $N.l\pi$ interaction as an example, we find that the π meson can be emitted also in the real state, and, in such a case, treatments disregarding the recoil effect are not justified.

The interaction of this type has been treated by Cheston and Primakoff¹²⁾ for the non-mesonic decay of the hyperfragment. They have calculated the transition from the Λ particle to the nucleon in a π meson field around another nucleon source. But they have neglected the transition of the nucleon in a π meson field around the Λ source. The latter contribution seems to be more efficient in this decay process. Moreover, the mass change in the case of the transition is a quantum effect, hence such a semi-classical treatment can not be applied. In other words, the static meson field does not relate

directly with the potential or the transition matrix element. About this point we shall discuss further in the Appendix.

The author wishes to express his cordial thanks to Prof. H. Umezawa and Dr. T. Okabayashi for their helpful discussions. He is also indebted to Prof. M. Nogami, Prof. K. Ono and many colleagues.

Appendix

We consider the static meson field and clarify its relation to the Equations (4), (4') derived by the calculation in the perturbation theory.

Let us take the equation of motion for the meson field variable φ_0 in the Heisenberg representation :

$$(\square - \mu^2)\varphi_0 = -g(\bar{\psi}_N\psi_\Lambda + \bar{\psi}_{\bar{N}}\psi_{\bar{N}}) \quad (\text{A} \cdot 1)$$

where

$$\bar{\psi}_N\psi_\Lambda = \frac{1}{(2\pi)^3} \int \bar{\psi}(k_0 - \delta) \exp(ik_\mu x_\mu) d^4k.$$

This source function means that the recoil effect of the fermion is neglected. The solution for the static field is given by

$$\begin{aligned} \varphi_0 &= \frac{g}{(2\pi)^3} P \int \frac{1}{k_\mu^2 + \mu^2} \bar{\psi}(k_0 - \delta) \exp(ik_\mu x_\mu) d^4k \\ &= \frac{g}{4\pi r} \exp(-\mu' r) \exp(-i\delta t). \end{aligned} \quad (\text{A} \cdot 2)$$

This is just the standing wave solution for Equation (A.1). The interaction between other fermion and this field is $-g\bar{\psi}_\Lambda\psi_N\varphi_0$. (Here the time-dependent factor $\exp(-i\delta t)$ does not give any effect on the potential or on the transition probability.) Therefore, the potential for NA system is given by

$$V = -\frac{g^2}{4\pi r} \exp(-\mu' r). \quad (\text{A} \cdot 3)$$

This is in agreement with the result of the perturbation theory; Equation (8).

But we have calculated here only the effect of the interaction of the nucleon with the boson field coming from the A -particle source. Thus, it may be expected that we should not obtain Equation (8), but Equation (4'). It is not apparent why such a semi-classical treatment has led to the potential (8) between two fermions of different masses, although this treatment has been justified in the cases of the Møller scattering of the two electrons and the potential of the nuclear force.

In order to clarify this, we investigate the solution (A.2). Introducing the following notations

$$\begin{aligned} I_1(x') &= -g\bar{\psi}_N(x')\psi_\Lambda(x'), \\ J_2(x'') &= -g\bar{\psi}_\Lambda(x'')\psi_N(x'') \\ &= -\frac{g}{(2\pi)^3} \int \bar{\psi}(k_0 + \delta) \exp(ik_\mu x''_\mu) d^4k, \end{aligned}$$

the solution (A·2) can be written as

$$\varphi_0(x) = - \int \bar{J}(x-x') J_1(x') d^4x'.$$

The interaction between this field and other fermion is given by

$$\begin{aligned} I &= - \int J_2(x'') \varphi_0(x'') d^4x'' \\ &= \iint J_2(x'') \bar{J}(x''-x') J_1(x') d^4x' d^4x''. \end{aligned} \quad (\text{A} \cdot 4)$$

If we use the following relations

$$\begin{aligned} \bar{J}(x) &= -(\varepsilon(x)/2) J(x), \quad J(x) = J^+(x) + J^-(x), \\ J^-(x) &= -J^+(-x), \quad J^{(1)}(x) = i(J^+(x) - J^-(x)), \end{aligned}$$

the integrand of I can be transformed into

$$\begin{aligned} &-J_2(x'') \frac{1+\varepsilon(x''-x')}{2} J^+(x''-x') J_1(x') \\ &-J_1(x') \frac{1+\varepsilon(x'-x'')}{2} J^+(x'-x'') J_2(x'') - \frac{i}{2} J_2(x'') J^{(1)}(x''-x') J_1(x'). \end{aligned}$$

Since the last term does not give any contribution to the result, we can write

$$\begin{aligned} I &= \iint J_2(x'') J_{\text{ret}}^+(x''-x') J_1(x') d^4x' d^4x'' \\ &+ \iint J_1(x') J_{\text{ret}}^+(x'-x'') J_2(x'') d^4x' d^4x''. \end{aligned} \quad (\text{A} \cdot 5)$$

This consists of the two terms, each of which represents the effect of the interactions between one fermion and a boson emitted in the past by another fermion. Thus, it is expected that Equations (4') derived from the perturbation theory are in agreement with the respective term in Equation (A·5). This is confirmed by the evaluation of Equation (A·5). The coincidence between Equation (A·3) and the result of the perturbation theory comes from the fact that the changes of masses of two sources are of the same values and of different signs.

Thus, the semi-classical treatment is not generally valid; it does not seem to be applicable, e.g., to the case treated by Cheston and Primakoff (cf. § 4). For, in such a case, the meson field is hardly to be the standing wave; the latter is composed of the outgoing and incoming waves of the same strength.

References

- 1) G. C. Wick, *Nature* **142** (1938), 993.
- 2) K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285.
- 3) M. Gell-Mann and A. Pais, *Proceedings of the Glasgow Conference* p. 342 (1954).
- 4) K. Nishijima, *Prog. Theor. Phys.* **14** (1955), 527.
- 5) S. Iwao, *Prog. Theor. Phys.* **13** (1955), 111.
- 6) W. Cheston and H. Primakoff, *Phys. Rev.* **92** (1953) 1537; **93** (1954), 908.

Free-Electron Ring Model of Polyacenes*

Gentaro ARAKI

Faculty of Engineering, Kyoto University, Yosida, Kyoto

(Received June 27, 1956)

Electronic energy levels of polyacenes are calculated by assuming their π electrons to move freely in a ring of a circumference which is equal to the perimeter length of the polyacene molecule. The interaction between π -electrons is assumed to be the long range part of the Coulomb interaction. The short range part is omitted by the zero cut-off at $0.48l$ of the arc length where l is the distance between adjacent carbons in the molecule. The agreement between calculated and observed levels is fairly well.

Introduction

The free-electron model of benzene was examined in the previous paper⁽²⁾ where the long range interaction between π -electrons was assumed. In the present paper the same idea is applied to polyacenes as the second step to examine the free-electron model for conjugated cyclic system. As was done by Platt⁽³⁾ the π -electrons are assumed to be enclosed in a ring of the circumference which is equal to the perimeter length of the polyacene molecule. It is further assumed that the interaction between π -electrons is given by the long range part of the Coulomb interaction. The significance of such a model was already accounted for in a previous paper⁽¹⁾. In what follows we shall see that this model is rather satisfactory in spite of its simpleness.

The electronic states of polyacenes were already discussed by other methods⁽⁵⁾. The free-electron model was also applied to this problem by taking into account the net-work structure of the molecule but not the interaction between π -electrons.⁽⁶⁾ Therefore the present model is another extreme as compared with the net-work model. It may be desired to take into account both the net-work structure and the interaction. However, this may be difficult and gives a too complicated aspect to a model theory for which the simpleness is characteristic. It was seen that the results of the free-electron net-work model were as good as those of the simplified LCAO calculations. We shall see that the present model gives the electronic levels which are as good as those given by the free-electron net-work model.

Calculated result and comparison with experiment

We consider the π -electron system of polyacene which consists of n benzene rings.

* A part of the preliminary result was already reported elsewhere⁽¹⁾.

The total number of π -electrons is equal to $2(2n+1)$ which is denoted by N . The perimeter length, L , of the polyacene molecule is equal to $2(2n+1)l$ where l is the bond length between two adjacent carbons. The Hamiltonian of the system is given by

$$H = - (1/2) \sum_{k=1}^N d^2/ds_k^2 + \sum_{k=1}^N \sum_{j=k+1}^N J(s_k - s_j) \quad (1)$$

where all quantities are measured in atomic units, s_k is the ring coordinate of the k -th electron, and $J(s_k - s_j)$ represents the interaction between two electrons. The ring coordinate

is measured by the arc length (including the sign) from a definite point. The interaction is defined by

$$J(s_i - s_j) = 0 \quad |s_k - s_j| < s_0 \\ = 1/r_{kj} \quad |s_k - s_j| \geq s_0 \quad (2)$$

where r_{kj} is the chord length corresponding to the arc given by $s_k - s_j$. This interaction is the long range part of the Coulomb interaction.

If we neglect the interaction the one-electron eigenfunction and the corresponding energy levels are given by

$$\varphi_q(s) = 1/\sqrt{L} \cdot e^{2\pi i q s/L} \\ \varepsilon(q) = 2(\pi/L)^2 q^2 \quad (3) \\ q = 0, \pm 1, \pm 2, \dots$$

where q is the one-electron ring quantum number. We adopt this



Naphthalene



Anthracene



Naphthacene



Pentacene

Fig. 1. Molecular structure of polyacenes

function as the molecular orbitals of polyacenes. In the ground state of the molecule all levels up to $q = \pm n$ are occupied by π -electrons. We consider the four excited states in which one electron is excited from the orbitals with the ring quantum number of n , $n-1$, $-n+1$, and $-n$ respectively to that of $n+1$. The ring quantum number, Q , of the total system is equal to 1, 2, $2n$, and $2n+1$ in these states respectively where Q is defined by $\sum q$. We prefer Platt's way of denoting states to the group-theoretical one because the former can discriminate these ground and four excited states by different symbols, namely, A, B, C, K and L respectively.

The method of calculating electronic levels was already accounted for in the case of benzene²¹. As was there mentioned, the states of different ring quantum numbers do not interact because of the ring symmetry of the present model and we can not discriminate

the a and b states of Platt, e.g., B_a and B_b . In order to remove such a degeneracy the model should have the same symmetry as the real molecule: for example, we should introduce any asymmetry into the form of the model or its skeleton potential.

The direct integrals are independent of orbitals because the electron density is uniform throughout the ring, and they give no influence to the excitation energy of the molecule. The other interaction integrals can all be expressed in terms of the following function:

$$D(q) = (2/L) [\log \cot \{ \pi s_0 / (2L) \} - \sum_{k=0}^{q-1} 2(2k+1)^{-1} \cos \{ (2k+1) \pi s_0 / L \}]. \quad (4)$$

If we neglect the configuration interaction the heights of lower excited levels from the ground level, 1A , are given by

$$\begin{aligned} E({}^{1,3}B) - E({}^1A) &= 2(\pi/L)^2 \cdot (2n+1) + (1 \pm 1)D(1) - D(2n+1) \\ E({}^{1,3}L) - E({}^1A) &= 2(\pi/L)^2 \cdot (2n+1) \pm D(2n+1) \\ E({}^{1,3}C) - E({}^1A) &= 2(\pi/L)^2 \cdot 4n + D(1) + (1 \pm 1)D(2) - D(2n) \\ &\quad - D(2n+1) \\ E({}^{1,3}K) - E({}^1A) &= 2(\pi/L)^2 \cdot 4n + D(1) \pm D(2n) - D(2n+1) \end{aligned} \quad (5)$$

where the upper signs correspond to the singlets and the lower signs to the triplets.

Table 1. Energy levels of polyacenes

The heights of levels from the ground level, 1A , are shown in eV.
(): estimation; []: absorption peak; n : number of benzene rings.

| | n | | 1B_a | 1B_b | 3B_a | 3B_b | 1C_a | 1C_b |
|-------------|-----|---------------|----------------|-----------|----------------|-----------|----------------|-----------|
| Naphthalene | 2 | obs. calc. | 7.31 5.79 | 5.58 | (6.11) 4.65 | (4.34) | — 7.43 | 6.32 |
| Anthracene | 3 | obs. calc. | 6.45 5.06 | 4.77 | (5.54) 3.32 | (3.72) | — 6.80 | 5.58 |
| Naphthacene | 4 | obs. calc. | 5.83 4.49 | 4.46 | (4.34) 2.58 | (3.50) | [6.62] 6.19 | 5.46 |
| Pentacene | 5 | obs. calc. | (5.33) 4.04 | 4.09 | (4.34) 2.11 | (3.01) | — 5.65 | — |
| | n | | 1L_a | 1L_b | 3L_a | 3L_b | 1K | 3K |
| Naphthalene | 2 | obs. calc. | 4.34 3.03 | 3.99 | 2.64 4.65 | (4.49) | — 6.55 | — 3.49 |
| Anthracene | 3 | obs. calc. | 3.31 2.16 | 3.47 | 1.82 3.32 | 4.09 | — 5.48 | — 6.82 |
| Naphthacene | 4 | obs. calc. | 2.60 1.63 | (3.16) | 1.24 2.58 | (3.66) | — 4.69 | — 5.70 |
| Pentacene | 5 | obs. calc. | 2.14 1.33 | 2.93 | 0.71 2.11 | (3.47) | — 4.10 | — 4.91 |

The numerical calculation is carried out for $s_0=0.68l$ and $l=1.4\text{\AA}$. The result is shown in Table 1 and Table 2 in which the observed values of excitation energy are also shown⁷⁾. We see that the agreement between calculated and observed values is fairly satisfactory for 1B and 1L but not for 3B and 1L . The configuration interaction may improve the agreement because we saw in the case of benzene²⁾ that the configuration interaction made levels higher except 1B . In the case of benzene the somewhat different value of the division constant was adopted, namely $s_0=0.64l$. The difference is due to the neglect of the configuration interaction in the present case. It should be noted here that it is essential for the present model to take into account the interaction between electrons in order to obtain the agreement between the calculated and observed values.

Table 2. Interaction integrals in eV

| <i>n</i> | <i>D</i> (1) | <i>D</i> (2) | <i>D</i> (3) | <i>D</i> (4) | <i>D</i> (5) | <i>D</i> (6) | <i>D</i> (7) | <i>D</i> (8) | <i>D</i> (9) | <i>D</i> (10) | <i>D</i> (11) |
|----------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------|---------------|
| 1 | -0.5463 | -1.6473 | -1.3622 | -0.5820 | | | | | | | |
| 2 | 0.5725 | -0.5267 | -0.9231 | -0.9673 | -0.8098 | | | | | | |
| 3 | 0.8734 | -0.0053 | -0.4300 | -0.6323 | -0.6963 | -0.6676 | -0.5769 | | | | |
| 4 | 0.9568 | 0.2428 | -0.1361 | -0.3563 | -0.4786 | -0.5331 | -0.5380 | -0.5063 | -0.4482 | | |
| 5 | 0.9665 | 0.3694 | 0.0387 | -0.1691 | -0.3024 | -0.3843 | -0.4280 | -0.4422 | -0.4334 | -0.4068 | -0.3666 |

Note added in proof. Recently N. S. Ham and K. Ruedenberg [J. Chem. Phys. 25 (1956), 13] devised a method of introducing the electronic interaction into the free-electron net-work model of polyacenes.

References

- 1) G. Araki and T. Murai, J. Chem. Phys. 22 (1954), 954.
- 2) G. Araki, Prog. Theor. Phys. 16 (1956) in press.
- 3) J. R. Platt, J. Chem. Phys. 17 (1949), 484.
- 4) G. Araki, Prog. Theor. Phys. 16 (1956) in press.
- 5) For example,
Th. Förster, Zeits. f. phys. Chem. B41 (1938), 287.
P. Daudel and R. Daudel, J. Chem. Phys. 16 (1948), 639.
C. A. Coulson, Proc. Phys. Soc. A 60 (1948), 257.
W. Moffitt, J. Chem. Phys. 22 (1954), 320, 1820.
- 6) Y. Ooshika, Fusseiron Kenkyu (mimeographed circular in Japanese), No. 29 (1950), p. 16.
H. H. Perkampus, Z. Naturforschg. 7a (1952), 594.
C. W. Scherr, J. Chem. Phys. 21 (1953), 1582.
N. S. Ham and K. Ruedenberg, Technical Report (Chicago), 1953-1954, part 1, p. 97.
- 7) H. B. Klevens and J. R. Platt, J. Chem. Phys. 17 (1949), 470.
J. R. Platt, J. Chem. Phys. 18 (1950), 1168.
C. W. Scherr, J. Chem. Phys. 21 (1953), 1582.
S. P. McGlynn, M. R. Padhye and M. Kasha, J. Chem. Phys. 23 (1955), 593.

On the Grand Canonical Distribution Method of Statistical Mechanics

Tsunenobu YAMAMOTO and Hirotugu MATSUDA

Quantum Chemistry Laboratory, Chemistry Department, Kyoto University, Kyoto

(Received June 9, 1956)

Usually, the grand canonical distribution method is derived only through plausibility arguments, and up to the present its accurate proof has never been made basing directly on ergodic theorem. In order to give the method a sound basis, we attempt to reformulate this in the case of ordinary classical liquids, and show that it is practically equivalent to both the microcanonical and the canonical distribution methods. As the result, we can safely use any one of the three methods of statistical mechanics, even for the purpose of investigating phase change.

§ 1. Introduction

The development of equilibrium statistical mechanics has made possible to include a wide range of applications. No one doubts that this chapter of physics has already realized a large number of fruitful results. This fact, however, does not necessarily mean that the formalism has left no room to be discussed. It is true that we have three kinds of prescription—microcanonical, canonical and grand canonical distribution methods, and that in many cases one may use one or another as he likes. But it is not self-evident that the three methods are practically equivalent to each other, and any one has not been able to decide whether it is true or not. We may here cite a typical example. Classical fluids have non-negative compressibilities. The problem is why they do. Its proof is given by Van Hove,¹⁾ who started from the canonical distribution. If one uses the grand canonical distribution instead of the canonical one, the proof is almost self-evident, or even trivial after Van Hove. For the grand canonical distribution method gives in itself the non-negative compressibility, whatever intermolecular forces may be, just as the specific heat of any system is always non-negative by the canonical distribution. Katsura calculated exactly the equation of states of "small system"²⁾ as well as Husimi-Temperley's model,³⁾ and concluded that the canonical and grand canonical methods are not always equivalent. If the three prescriptions are not always equivalent to each other, or if they are not evenly reliable, it would not be a matter of small significance. We believe that such unpleasant circumstances are due solely to the fact that the standard methods of statistical mechanics have not been established on their solid bases. In other words, there still remains some gap between the ergodic theorem and the usual prescriptions of statistical mechanics. The purpose of the present article is to investigate the validity of the grand canonical distribution method in the case of ordinary classical liquids and to show that this method is practically equivalent to the other two, to the result that we can use safely any of the three for the above liquids.

The most reasonable formulation of statistical mechanics is, we think, what is given by Khinchin in his book.¹⁾ Although this book, as its title shows, lays emphasis on the mathematical frame of statistical mechanics, it also throws light upon its physical side and discloses various weak points of the theory. One of the characteristics of his formulation consists in separating the mathematical construction from the physical base. Simply speaking, the physical base is the supposition that the statistical mechanical system is composed of a very large number of mutually almost independent "components". The mathematical construction is nothing but the application of probability theory: the equilibrium properties of any system composed of a very large number of components can be asymptotically determined by the action of some central limit theorem.

At the same time, the recent development of the statistical or statistical mechanical theories of irreversible processes seems to suggest strongly that a statistical mechanical system is composed of a very large number of components which are subject to a kind of central limit theorem. For instance, Onsager and Machlup⁵ and Hashitsume⁶⁾ have formulated irreversible processes as Gaussian Markoff processes which could most probably be considered as asymptotic behaviors of groups of quite many components. Furthermore, when M. S. Green⁷⁾ and one of the present authors⁸ derived the above Markoffian processes from the first principle of mechanics, they laid stress on the assumption that gross variables are approximate one-valued integrals of motion, which, we believe, would be essentially equivalent to the concept of components. Thus we may say that in non-equilibrium as well as equilibrium states the concept of "component" is playing an extremely essential role to the result they can be treated statistically.

Before giving a brief outline of the present article, we should note on its limit of application. It refers to the nature of interparticle interactions and is expressed by the following conditions:

- i. The particles interact with each other only through two-body forces, the total potential energy being written as $U_N = \sum_{i>j=1}^N u(x_i, x_j)$.
- ii. The mutual potential energy $u(x_i, x_j)$ between the i -th and j -th particle has the properties:

$$(1) \quad u(x_i, x_j) = +\infty \quad \text{when} \quad |x_i - x_j| \leq d_0,$$

$$(2) \quad |u(x_i, x_j)| \leq \bar{u} \quad \text{when} \quad d_0 < |x_i - x_j| \leq d_1,$$

$$(3) \quad u(x_i, x_j) = 0 \quad \text{when} \quad d_1 < |x_i - x_j|,$$

where d_0 , d_1 and \bar{u} are constants.

In Sec. 2 we shall emboss the mathematical frame of statistical mechanics of liquids, making clear the action of probability theory in it. For this purpose we introduce an imaginary system which is defined rather *ad hoc* and which is called π -system. π -system consists, by definition, of a large number of components, so that Khinchin's probability-theoretical treatment is straight-forwardly applicable to it. In particular the above conditions allow us to use central limit theorem and lead to the usual grand canonical distribution.

In Sec. 3 a real system will be considered in relation to its π -system. When it is of a macroscopic size, the two systems are shown to have their common statistical mechanical properties. Here again the above nature of interparticle interactions plays an important role. Thus, π -system is supposed as a good representation of the real system and the grand canonical distribution method becomes a well established means for statistical mechanical computations. Next, this method is directly connected with thermodynamics. Properly speaking, the hydrostatic pressure should be derived by differentiating the logarithm of grand partition function with respect to volume. But one usually finds it by division instead of differentiation. Its reason is given there for the first time. In Sec. 4 the three methods - microcanonical, canonical and grand canonical distributions are shown practically equivalent to each other. The last section involves some comments on Katsura's works and a discussion on the occurrence of phase change in the frame of our theory.

§ 2. π -system

To simplify the discussion, we consider an isolated system composed of N identical particles and occupying a domain Γ . Suppose Γ is divided into M subdomains $\gamma_1, \gamma_2, \dots, \gamma_M$, all of which are identical cubic cells except for those involving the boundary of Γ .

Now we imagine a new liquid in which all the inter-cell interactions vanish. The particles are still allowed to move freely from cell to cell and within every cell they interact with each other through the same forces as in the real liquid. Since, in the case of real liquids, such neglects can not be done without special reasons, the now defined liquid is merely an imaginary one, which we call π -system. In the rest of this section we shall be concerned only with the statistical mechanical properties of this system and its relation to the real one will be given in the subsequent sections.

Due to ergodic theorem, the equilibrium properties of an isolated system with a constant energy E are determined by the microcanonical distribution. Now the knowledges obtainable by the usual methods of thermodynamic measurements are quite limited, that, from the theoretical point of view, it is sufficient to know from the above distribution the probability for one of the components composing the system to have its energy e and its number of particles n . For this purpose, as Khinchin made clear, we must investigate the asymptotic form of the structure function of π -system.

By definition, π -system is composed of components. One can, therefore, easily proceed along much the same line of reasoning as Khinchin's. Let $\tilde{\Omega}(E, N; \Gamma)$ be the structure function of π -system where E and N stand for the total energy and the total number of particles of the system respectively.* Now we introduce the two-dimensional Laplace transform $\tilde{\Xi}(\alpha, \lambda; \Gamma)$ of the structure function:

$$\tilde{\Xi}(\alpha, \lambda; \Gamma) = \sum_{N=1}^{\infty} \int_{-\infty}^{+\infty} e^{-\alpha E - \lambda N} \tilde{\Omega}(E, N; \Gamma) \frac{dE}{N! h^{3N}}, \quad (2.1)$$

* All the quantities with the symbol \sim refer to π -system, while those without \sim are reserved for the real system.

which assumes finite values^{*)} if

$$\alpha > 0, \quad -\infty < \lambda < +\infty. \quad (2.2)$$

Now, Khinchin's conjugate distribution may be written down as follows:

$$U(\alpha, \lambda; X, Y; I') = \frac{1}{\tilde{\Xi}(\alpha, \lambda; I')} e^{-\alpha X - \lambda Y} \tilde{Q}(X, Y; I') \frac{1}{Y! b^{3Y}}. \quad (2.3)$$

As is well known, the means and variances with respect to the conjugate distribution can be calculated by differentiating the logarithm of its generating function $\tilde{\Xi}(\alpha, \lambda; I')$:

$$\langle X \rangle = -\partial \log \tilde{\Xi} / \partial \alpha, \quad \langle Y \rangle = -\partial \log \tilde{\Xi} / \partial \lambda, \quad (2.4)$$

$$D_{EE} \equiv \langle (X - \langle X \rangle)^2 \rangle = \partial^2 \log \tilde{\Xi} / \partial \alpha^2, \quad D_{NN} \equiv \langle (Y - \langle Y \rangle)^2 \rangle = \partial^2 \log \tilde{\Xi} / \partial \lambda^2,$$

$$D_{EN} \equiv \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle = \partial^2 \log \tilde{\Xi} / \partial \alpha \partial \lambda. \quad (2.5)$$

Corresponding to the fact that each cell acts as a component of π -system, the following probability-theoretical relation can be easily obtained:

$$U(\alpha, \lambda; X, Y; I') = \sum_{\Sigma y_i = Y} \dots \sum_{\Sigma x_i = X} \int \dots \int \prod_{i=1}^M U(\alpha, \lambda; x_i, y_i; \gamma_i) dx_i. \quad (2.6)$$

From Eq. (2.6) we can see at once that the two-dimensional random variable (X, Y) is just equal to the sum of M independent random variables (x_i, y_i) . Accordingly, if M is sufficiently large and if the two-dimensional central limit theorem is applicable to Eq. (2.6), the asymptotic form of $U(\alpha, \lambda; X, Y; I')$ is given as follows:

$$U(\alpha, \lambda; X, Y; I') = \frac{1}{2\pi J} \exp \left[-\frac{1}{2J^2} \{ D_{NN}(X - \langle X \rangle)^2 + \right. \\ \left. + D_{EE}(Y - \langle Y \rangle)^2 - 2D_{EN}(X - \langle X \rangle)(Y - \langle Y \rangle) \} \right], \quad (2.7)$$

where

$$J^2 = D_{EE} D_{NN} - D_{EN}^2. \quad (2.8)$$

Now we must make a remark as regards the applicability of the limit theorem to our problem. In view of the assumed nature of the inter-particle interaction, it seems almost self-evident that this theorem holds in our case. Namely any cell can involve not more than a finite number of particles and, as a result, a finite value of potential energy. $U(\alpha, \lambda; x, y; \gamma)$ has, therefore, its finite absolute moments of lower orders.

The asymptotic $U(\alpha, \lambda; X, Y; I')$ given by Eq. (2.7) still contains two undetermined parameters α and λ . If we put

$$\langle X \rangle = E, \quad \langle Y \rangle = N, \quad (2.9)$$

* Because of the hard core potential \sum_N in (2.1) is practically a finite series. Thus the variable E has its finite lower limit. Besides $\tilde{Q}(E, N; I')$ can be assumed to increase slowly compared with $e^{-\alpha E}$. As the result, we can safely assume the existence of $\tilde{\Xi}(\alpha, \lambda; I')$ for any real finite values of $\alpha > 0$ and λ .

where E and N refer to the isolated π -system under consideration, these equations uniquely fix the values of α and λ , which we shall hereafter denote by β and μ (see App. 3). Then the asymptotic form of the structure function is given by

$$\tilde{Q}(E, N; \Gamma) = (2\pi\Delta)^{-1} \tilde{\Xi}(\beta, \mu; \Gamma) e^{\beta E + \mu N} (1 + o(1)). \quad (2 \cdot 10)$$

Having obtained the asymptotic form of the structure function, it would no longer be necessary to present the straightforward calculations for finding the probability distribution $p(e, n; \gamma)$ of a single component. The result is the wellknown grand canonical distribution:

$$p(e, n; \gamma) = (1/\tilde{\Xi}(\beta, \mu; \gamma)) e^{-\beta e - \mu n} \tilde{Q}(e, n; \gamma). \quad (2 \cdot 11)$$

Looking upon e as a function of external parameters, differentiating it with respect to these parameters, and using Eq. (2.11) to get an average, one can calculate the work done by the external forces. Thus, for example, the hydrostatic pressure \tilde{P} is formally derived in the form:

$$\beta \tilde{P} = (\partial/\partial V) \log \tilde{\Xi}(\beta, \mu; \Gamma), \quad (2 \cdot 12)$$

where V stands for the volume of Γ .

Since no discussions have so far been made on the size of cell except that it is finite, we are obliged to suppose the pressure given by Eq. (2.12) to depend not only on β and μ but also on cell size. Besides, it should be noted that Eq. (2.12) is apparently different from the usual expression that is known in statistical thermodynamics:

$$\beta P = (1/V) \log \Xi(\beta, \mu; I'). \quad (2 \cdot 13)$$

Indeed, as far as we remain with π -system, neither the equivalence between the two pressures nor their cell size-independence can be shown. Anyway, π -system is, by definition, an imaginary one, and we need not discuss it any further. In the following sections we shall go back to the real system and investigate its statistical mechanical properties, where π -system will most effectively serve as a mathematical device, the results of the present section being fully utilized.

§ 3. Grand canonical distribution of a cell

In π -system the probability density for a cell γ_1 to have the potential energy $e_1^{(*)}$ and the number of particles n_1 , when the total potential energy and the total number of particles of the whole system I' are E and N , is given according to the microcanonical distribution by

$$\tilde{Q}(e_1, n_1; \gamma_1) \tilde{Q}(E - e_1, N - n_1; \Gamma - \gamma_1) / \tilde{Q}(E, N; \Gamma). \quad (3 \cdot 1)$$

Making use of central limit theorem, this has been shown in Sec. 2 to become the well-

* For the sake of simplicity we omit the kinetic energy, since its inclusion offers no new difficulties in the following discussions.

known grand canonical distribution law provided that the number of cells M is sufficiently large.

In the real system, the potential energy of γ_1 is composed of two parts: the *intracellular* potential e_1 , and the *intercellular* one e_{12} which is completely omitted in π -system. We must find out the distribution law of e_1 or at least the mean value of e_1 , for this will become indispensable when we shall eventually make a comparison with thermodynamics.

As is shown in Appendix 1, there exists a remarkable tendency that π -system asymptotically simulates the real one. It is proved there that the logarithm of the grand partition function of the real system is asymptotically equal to that of the corresponding π -system. Nevertheless, we are not yet free, in deriving the distribution law mentioned above, to substitute π -system for the real one from the beginning. Indeed, for the real system the situation seems troublesome *prima facie*. Because of the intercellular interaction we cannot compose structure functions by convolution, so we cannot at once expect that the distribution law can be given in such a simple form as Eq. (3.1). A little consideration, however, will suffice to show that we can proceed quite analogously as in the case of π -system *mutatis mutandis*.

We construct in each cube γ a cube γ' with the same centre and the faces parallel to γ , the side length of which is $d-2d_1$, where d is the side length of the original cubic cell and d_1 is the range of potential defined in Sec. 1. Given the domain I' composed of cubes $\gamma_1, \gamma_2, \dots, \gamma_M$, we denote by I'' the domain composed of the cubes $\gamma'_1, \gamma'_2, \dots, \gamma'_M$, and by $I-I''$ the part of I not belonging to I'' .

If we fix all the positions $\{X\}$ of particles that are involved in $I-I''$, the total intercellular interaction energy is determined, and these fixed particles in $I-I''$ can be looked upon as the external sources of force acting on the moving particles in I'' . Noting these circumstances, we can now ask for the probability density that the intracellular potential of γ_1 is e_1 and the number of particles within γ'_1 (not γ_1) is n'_1 when we completely know the configuration $\{X\}$ of the particles in $I-I''$. (conf. Fig. 1).

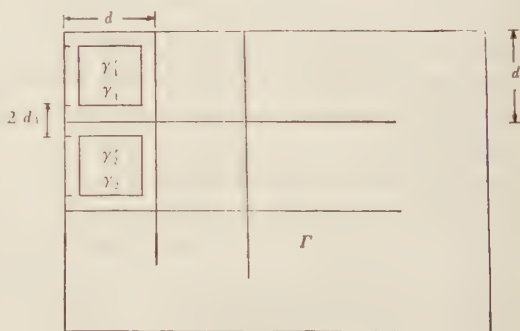


Fig. 1.

Owing to the assumed nature of inter-particle interaction, γ'_1 has no direct interaction with $I-\gamma_1$. Looking upon $\{X\}$ as parameters, therefore, the answer is

$$p(e_1, n'_1; \{X\}) = \frac{\mathcal{Q}(e_1, n'_1; \{X\}, \gamma_1) \mathcal{Q}(E-E_1-e_1, N-N_1-n'_1; \{X\}, I-\gamma_1)}{\mathcal{Q}(E-E_1, N-N_1; \{X\}, I)} \quad (3.2)$$

Here $\mathcal{Q}(e_1, n'_1; \{X\}, \gamma)$ is the structure function of γ' in the static field due to $\{X\}$

While e_1 is supposed to include all the intracellular potential energy in γ , n' refers to the particle number in γ' . N_1 is the particle number in $\Gamma - \Gamma'$ and E_1 is the total intercellular potential energy between the N_1 particles.*)

An important conclusion derivable from Eq. (3.2) is that, in the situation where the positions of all the particles in $\Gamma - \Gamma'$ are fixed, Γ can be considered as a π -system under the influence of external force field. Therefore, it becomes again possible to apply central limit theorem. The result is, when M is sufficiently large,

$$p(e_1, n_1'; \{X\}) = \mathcal{Q}(e_1, n_1'; \{X\}, \gamma_1) \frac{e^{-\beta' e_1 - \mu' n_1'}}{\Xi(\beta', \mu'; \{X\}, \gamma_1)} \{1 + o(1)\}, \quad (3.3)$$

where

$$\Xi(\alpha, \lambda; \{X\}, \gamma) = \sum_{n=1}^{\infty} \int e^{-\alpha e - \lambda n} \mathcal{Q}(e, n; \{X\}, \gamma) \frac{de}{n! h^{3n}}, \quad (3.4)$$

and β' and μ' are determined by

$$-\partial \log \Xi(\beta', \mu'; \{X\}, \Gamma) / \partial \beta' = E - E_1, \quad -\partial \log \Xi(\beta', \mu'; \{X\}, \Gamma) / \partial \mu' = N - N_1. \quad (3.5)$$

Using Eq. (3.3), we can calculate the mean value of e_1 when the set $\{X\}$ is prescribed:

$$\langle e_1(\{X\}, \gamma_1) \rangle = -\partial \log \Xi(\beta', \mu'; \{X\}, \gamma_1) / \partial \beta'. \quad (3.6)$$

When e_1 is the function of extensive parameters $\lambda_1, \lambda_2, \dots, \lambda_r$ the generalized external force $\langle A_s \rangle$ in the direction of λ_s is

$$\begin{aligned} \langle A_s(\{X\}, \gamma_1) \rangle &= \sum_{n_1'} \int \frac{\partial e_1}{\partial \lambda_s} \mathcal{Q}(e_1, n_1'; \{X\}, \gamma_1) \frac{e^{-\beta' e_1 - \mu' n_1'}}{\Xi(\beta', \mu'; \{X\}, \gamma_1)} de_1 \\ &= -(1/\beta') \partial \log \Xi(\beta', \mu'; \{X\}, \gamma_1) / \partial \lambda_s. \end{aligned} \quad (3.7)$$

What we want to know is $\langle A_s(\gamma_1) \rangle$ which is the generalized force when the set $\{X\}$ is not prescribed or when $\langle A_s(\{X\}, \gamma_1) \rangle$ is averaged over all the possible values of the set $\{X\}$. That is to say, our aim is to show that if the sizes of the cell γ_1 and of Γ are sufficiently large, $\langle A_s(\gamma_1) \rangle$ is given by

$$\langle A_s(\gamma_1) \rangle = -(1/\beta) \partial \log \Xi(\beta, \mu; \gamma_1) / \partial \lambda_s, \quad (3.8)$$

where β and μ are determined by

$$-\partial \log \Xi(\beta, \mu; \Gamma) / \partial \beta = E, \quad (3.9)$$

and

$$-\partial \log \Xi(\beta, \mu; \Gamma) / \partial \mu = N.$$

* e_1 is the sum of potential energy acting between the fixed particles, between the unfixed ones and between the fixed and unfixed particles in γ_1 . Intercellular potential energy is the sum of potential energy acting between a pair of particles occupying different cubes among $\gamma_1, \gamma_2, \dots, \gamma_M$. E_1 is solely due to the N_1 fixed particles. Letting $e(\gamma_i)$ be the intracellular potential energy of γ_i , we have the relation

$$\sum_{i=1}^M e(\gamma_i) + E_1 = E.$$

In order to prove this, it will suffice to show that in the limit of infinite cell size,

(I) the values of β' and μ' determined by (3.5) are equal to the values of β and μ determined by (3.9); and that

(II) when the uniformity of convergence of the two limiting processes (3.24) hold,

$$\partial \log \Xi(\beta, \mu; \{X\}, \gamma_1) / \partial \lambda_1 \quad \text{is equal to} \quad \partial \log \Xi(\beta, \mu; \gamma_1) / \partial \lambda_1. *)$$

Putting $\Xi(\beta, \mu; \Gamma) = \exp[V(\Gamma)g(\beta, \mu; \Gamma)]$ etc. (cf. App. 1), Eqs. (3.5) and (3.9) can be rewritten as follows:

$$-\partial g(\beta', \mu'; \{X\}, \Gamma) / \partial \beta' = (E - E_1) / V(\Gamma), \quad (3.5')$$

$$-\partial g(\beta', \mu'; \{X\}, \Gamma) / \partial \mu' = (N - N_1) / V(\Gamma), \quad (3.5'')$$

$$-\partial g(\beta, \mu; \Gamma) / \partial \beta = E / V(\Gamma), \quad (3.9')$$

$$-\partial g(\beta, \mu; \Gamma) / \partial \mu = N / V(\Gamma). \quad (3.9'')$$

We notice that

$$\begin{aligned} \lim_{\substack{M \rightarrow \infty \\ d \rightarrow \infty}} (N - N_1) / V(\Gamma) &= \lim_{\Gamma(\Gamma) \rightarrow \infty} N / V(\Gamma), \\ \lim_{\substack{M \rightarrow \infty \\ d \rightarrow \infty}} (E - E_1) / V(\Gamma) &= \lim_{\Gamma(\Gamma) \rightarrow \infty} E / V(\Gamma), \end{aligned} \quad (3.10)$$

which can easily be verified from the inequalities:

$$N_1 \lesssim Md_1 d^2 / v_0, \quad |E_1| \leq \bar{c} u N_1. \quad (3.11)$$

Here, $v_0 = 4\pi/3 \cdot (d_0/2)^3$ and c is a number of order unity. These are the direct consequences from the assumptions imposed on the inter-particle interaction.

In App. 1, it is proved that

$$\lim_{\substack{d \rightarrow \infty \\ \Gamma(\Gamma) \rightarrow \infty}} g(\beta, \mu; \{X\}, \Gamma) = \lim_{\Gamma(\Gamma) \rightarrow \infty} g(\beta, \mu; \Gamma) = \mathcal{T}(\beta, \mu). \quad (3.12)$$

The convergence is shown there to be uniform with respect to both β and μ when β is bounded. As a result, the limit function $\mathcal{T}(\beta, \mu)$ is a continuous function of β and μ .

On the other hand, from the definition it is easy to see that

$$\partial^2 g(\beta, \mu; \{X\}, \Gamma) / \partial \beta^2 > 0, \quad \partial^2 g(\beta, \mu; \{X\}, \Gamma) / \partial \mu^2 > 0, \quad (3.13)$$

and

$$\partial^2 g(\beta, \mu; \Gamma) / \partial \beta^2 > 0, \quad \partial^2 g(\beta, \mu; \Gamma) / \partial \mu^2 > 0. \quad (3.14)$$

Let $\beta_1, \beta_1', \beta_2$ and β_2' be any numbers such that

$$\begin{aligned} \beta_1' < \beta_1 < \beta < \beta_2 < \beta_2' \\ \beta_1 < \beta' < \beta_2. \end{aligned} \quad (3.15)$$

* When we make $V(\Gamma)$ or d tend to infinity, we suppose $\{X\}$ to be an arbitrary function of $V(\Gamma)$ and d . The equality in the limit implies that it holds irrespective of the form of the arbitrary function.

From (3.13) and (3.14) we have

$$\begin{aligned} \frac{g(\beta_1, \mu'; \{X\}, \Gamma) - g(\beta_1', \mu'; \{X\}, \Gamma)}{\beta_1 - \beta_1'} &< \frac{\partial g(\beta', \mu'; \{X\}, \Gamma)}{\partial \beta'} \\ &< \frac{g(\beta_2', \mu'; \{X\}, \Gamma) - g(\beta_2, \mu'; \{X\}, \Gamma)}{\beta_2' - \beta_2}, \end{aligned} \quad (3.16)$$

$$\frac{g(\beta_1, \mu; \Gamma) - g(\beta_1', \mu; \Gamma)}{\beta_1 - \beta_1'} < \frac{\partial g(\beta, \mu; \Gamma)}{\partial \beta} < \frac{g(\beta_2', \mu; \Gamma) - g(\beta_2, \mu; \Gamma)}{\beta_2' - \beta_2}. \quad (3.17)$$

Let

$$L(\beta_1, \beta_2, \mu) \equiv \lim_{\substack{d \rightarrow \infty \\ V(\Gamma) \rightarrow \infty}} \frac{g(\beta_1, \mu; \Gamma) - g(\beta_2, \mu; \Gamma)}{\beta_1 - \beta_2}. \quad (3.18)$$

Then, using (3.12) and continuity of $\tilde{g}(\beta, \mu)$, we have

$$L(\beta_1, \beta_1', \mu) + \varepsilon_1 + \varepsilon_2 < \partial g(\beta', \mu'; \{X\}, \Gamma) / \partial \beta' < L(\beta_2', \beta_2, \mu) + \varepsilon_1' + \varepsilon_2', \quad (3.19)$$

$$L(\beta_1, \beta_1', \mu) + \varepsilon_1'' < \partial g(\beta, \mu; \Gamma) / \partial \beta < L(\beta_2', \beta_2, \mu) + \varepsilon_1''', \quad (3.20)$$

where ε_1 , ε_1' , ε_1'' and ε_1''' tend to zero as both d and $V(\Gamma)$ tend to infinity, and

$$\lim_{|\mu - \mu'| \rightarrow 0} \varepsilon_2 = 0, \quad \lim_{|\mu - \mu'| \rightarrow 0} \varepsilon_2' = 0. \quad (3.21)$$

Using (3.5) and (3.10) we can rewrite (3.20) in the form

$$L(\beta_1, \beta_1', \mu) + \varepsilon_1'' < -(E - E_1) / V(\Gamma) < L(\beta_2', \beta_2, \mu) + \varepsilon_1'''. \quad (3.22)$$

Comparing (3.19) with (3.22) and noting (3.13), we can conclude in the case of sufficiently large $V(\Gamma)$ that, if the value of μ' is sufficiently near to that of μ , there is a value of β' for any μ' which satisfies (3.5) and the relation:

$$|\beta - \beta'| < \beta_2' - \beta_1'. \quad (3.23)$$

Since we can choose $|\beta_2' - \beta_1'|$ as small as we wish, if we take μ' sufficiently near to μ there is the value of β' which is sufficiently near to β and which satisfies Eq. (3.5'). The same can be said with Eq. (3.5'').

Thus, there is the set of values of β' and μ' satisfying Eqs. (3.5') and (3.5'') which tend to those of β and μ in the limit. As there is the proof for the uniqueness of the solution of (3.5'), (3.5''), (3.9'), and (3.9'') in App. 3, we can make the above assertion (I). Although we have the relations (3.12), they are not sufficient for the proof of (II).

If we could ascertain the uniformity of convergence of the two limiting processes:

$$\lim_{\substack{d \rightarrow \infty \\ V(\Gamma) \rightarrow \infty}} \partial \log \Xi(\beta, \mu; \{X\}, \gamma_1) / \partial \lambda_s \quad \text{and} \quad \lim_{V(\Gamma) \rightarrow \infty} \partial \log \Xi(\beta, \mu; \Gamma) / \partial \lambda_s, \quad (3.24)$$

there would be no difficulty to prove (II).

Unfortunately, we have not succeeded to attain the rigorous proof. Since it is a rather troublesome task to show strictly the above uniformity of convergence, we have given up this problem. From the physical point of view, however, it will be quite natural

to expect (II) because we have the relations (3.12).*) Incidentally we may remark that this point remains untouched in the theory of Van Hove¹⁾ as well.

The conclusion we have just reached is that the grand canonical distribution law is under certain conditions effectively applicable to the real system, as if the intercellular interactions were infinitely small as long as both d and M are sufficiently large. In short, our π -system is a very good replica of the real system in the statistical mechanical sense.

At the end of this paragraph, a word must be added as to the equivalence between (2.12) and (2.13). As is fully discussed in App. 1, $\log \Xi(\beta, \mu; \Gamma)$ is asymptotically equal to $V(\Gamma) \tilde{\eta}(\beta, \mu)$ where $\tilde{\eta}(\beta, \mu)$ is independent of $V(\Gamma)$. Thus the above equivalence is now obvious. As is easily recognized, this fact plays an essential role in deriving thermodynamics from our formulation. Since the derivation is straightforward, it may be omitted here.

§ 4. The equivalence of various formalisms of statistical mechanics.

It is well-known that there are various formalisms of statistical mechanics. They are different in defining the entropy of a closed system Γ as a function of the number of particles N and the total energy E of the system. Namely:

(1) Boltzmann's method:

$$S_B(E, N; \Gamma) = k \log \{V_N(E + \Delta E; \Gamma) - V_N(E; \Gamma)\}, \quad (4.1)$$

where ΔE is a constant independent of N and E , k is the Boltzmann constant and $V_N(E; \Gamma)$ is the phase volume of the set of points the total energy corresponding to which is smaller than E .

(2) Method of canonical distribution:

$$S_C(E, N; \Gamma) = k \{ \log \Phi(\beta, N; \Gamma) + \beta E \}, \quad (4.2)$$

where β is determined in terms of the ordinary partition function $\Phi(\beta, N; \Gamma)$ by the relation:

$$-\partial \log \Phi / \partial \beta = E. \quad (4.3)$$

(3) Method of grand canonical distribution:

$$S_G(E, N; \Gamma) = k \{ \log \Xi(\beta, \mu; \Gamma) + \beta E + \mu N \}, \quad (4.4)$$

where β and μ are determined by Eq. (3.9).

Of these three methods we have seen in the preceding section that the method of grand canonical distribution is well founded in the sense that by this method the thermodynamic properties of the system under consideration are obtainable on the basis of the ergodic theorem.

As to the other methods plausibility arguments for their validity have been given

* Practically speaking, we should take difference quotient rather than differential one with respect to λ_s . From this point of view the assertion (II) is now obvious from (3.12).

in various textbooks, and yet the rigorous proof for the system of dependent particles does not seem to exist. The equivalence of these methods will not be so self-evident especially when the system is undergoing a phase change.

The results we have obtained in this respect are the following: There exist

$$\lim_{N \rightarrow \infty} S_R(E, N; \Gamma) / N = s_R(\varepsilon, \rho), \quad (4.5)$$

$$\lim_{N \rightarrow \infty} S_C(E, N; \Gamma) / N = s_C(\varepsilon, \rho), \quad (4.6)$$

$$\lim_{N \rightarrow \infty} S_G(E, N; \Gamma) / N = s_G(\varepsilon, \rho), \quad (4.7)$$

and

$$s_R(\varepsilon, \rho) = s_C(\varepsilon, \rho) = s_G(\varepsilon, \rho), \quad (4.8)$$

where in the limiting process $N/V(\Gamma)$ and $E/V(\Gamma)$ are fixed at ρ and ε respectively.

Making use of π -system, these results can be directly obtained from the following lemmas. Here, we use the symbol \tilde{X}^d for the quantity X of π -system whose basic cube has the side length d .

$L(2.1)$: There exists

$$\tilde{s}_G^d = \lim_{N \rightarrow \infty} \tilde{S}_G^d / N, \quad (4.9)$$

and it is a function of ρ and ε only.

$L(2.2)$: When M is large,

$$\begin{aligned} \tilde{S}_R^d(E, N; \Gamma) &= \tilde{S}_C^d(E, N; \Gamma) + o(M) \\ &= \tilde{S}_G^d(E, N; \Gamma) + o(M). \end{aligned} \quad (4.10)$$

$L(2.3)$: When d is large,

$$\tilde{s}_R^d(\varepsilon, \rho) = s_R(\varepsilon, \rho) + O(1/d). \quad (4.11)$$

$L(2.4)$: When d is large,

$$\tilde{s}_C^d(\varepsilon, \rho) = s_C(\varepsilon, \rho) + O(1/d), \quad (4.12)$$

$$\tilde{s}_G^d(\varepsilon, \rho) = s_G(\varepsilon, \rho) + O(1/d). \quad (4.13)$$

The proofs of all the statements of this section are given in the appendix 2.

It is to be noted that the above results are quite independent of the existence of phase change.

§ 5. Concluding Remarks

When a system is of a macroscopic size and when the interparticle interaction satisfies the conditions given in Sec. 1, the grand canonical distribution method has been exactly founded on the basis of ergodic theorem. Then the equilibrium properties of the system have been shown to be equally derivable from anyone of the usual statistical mechanical methods—microcanonical, canonical, and grand canonical methods.

As to the assumptions imposed on the nature of the interparticle interaction, it can be shown that they are unnecessarily restrictive. One may allow not only two-body force but also n -body force so long as n is finite and independent of the total number of particles. Further, the two-body force does not necessarily have a hard core and it is the point that the force increases with sufficient rapidity when two particles approach to each other nearer and nearer. Anyway, our proof can be supposed to apply to the ordinary classical liquids and gases.

Katsura^{21,3)} attempted to compare the results of canonical and grand canonical distribution methods with each other. Since exact calculations must be carried out to eliminate any ambiguities, he had to refer to rather *ad hoc* models; i.e., "small clusters" and the Husimi-Temperley model. Now, it is not surprising that his results show in both cases decisive discrepancies between the two methods, because on the one hand his small cluster is of a *microscopic size* and because on the other hand the Husimi-Temperley model involves many-body forces which depend on the total number of particles of the system. This contradicts the assumptions made in the present theory and does not allow to go to the limit $d \rightarrow \infty$ with a constant particle density and energy density. Thus we can see that there are no reasons why the canonical and grand canonical distribution methods should give the same results in each of these models.

Finally it should be emphasized that the present theory holds equally well even when phase change is proceeding. In the present theory we have occasionally utilized the regular character of $\Xi(\beta, \mu; l')$ with respect to β and μ , which itself is a direct consequence of the assumed nature of inter-particle interaction. Now it might look paradoxical that a regular $\Xi(\beta, \mu; l')$ can produce discontinuous behaviors of a thermodynamic system. This point can be seen as follows: Let n_0 be the maximum of the number of particles that a cubic cell can contain. Various means and variances with respect to the conjugate distribution $U(\beta, \mu; X, Y)$ are of the order of magnitude $O(n_0)$ and $O(n_0^2)$ respectively. A quantity of $O(n_0^2)$ occasionally appears infinitely large compared with another quantity of $O(n_0)$ when n_0 is quite large. This means that an infinite heat capacity or a zero compressibility are practically possible, in spite of the regular character of $\Xi(\beta, \mu; l')$. Besides, however rapidly these variances may change during the course of changing β or μ , it does not effect the validity of the central limit theorem, so far as these variances do not exceed a finite value independent of the number of cubic cells M . Thus the conditions for the applicability of limit theorem do not exclude the possible occurrence of thermodynamic discontinuities.

Appendix 1

The technique of making use of π -system as a mathematical device to discuss the nature of the original system is nothing but the method which Van Hove¹⁾ applied ingeniously. The results he has derived, as to the relationship between the two systems, however, are concerned with canonical distribution, so it is necessary for our purposes to generalize them to the case of grand canonical formulation. The extension is straightforward. The reader is recommended to refer to Van Hove's paper.

We imagine that a cubic lattice L , with lattice constant d larger than d_0 is constructed in the space including the domain I' which contains the N particles in question, the space being divided by a set of cubes with side length equal to d .

Let I_1 be the maximum domain formed by the cubes completely contained in I' and I_2 be the minimum domain formed by the cubes covering I' .

We suppose that

(1) the volume of I' , $V(I')$, increases indefinitely with N in such a way that the density of the system $N/V(I')$ is fixed at a finite value ρ , and that

$$(2) \quad \lim_{V(I') \rightarrow \infty} V(I_1)/V(I_2) = 1. \quad (\text{A1} \cdot 1)$$

We denote by $\Phi(\beta, N; I')$ and $\Xi(\beta, \mu; I')$ the partition function and the grand partition function of the system under consideration, respectively.

Putting

$$\Phi(\beta, N; I') = \exp[Nf(\beta, N; I')], \quad (\text{A1} \cdot 2)$$

and

$$\Xi(\beta, \mu; I') = \exp[V(I')g(\beta, \mu; I')], \quad (\text{A1} \cdot 3)$$

it is easy to see that

$$f(\beta, N; I_1) \leq f(\beta, N; I') \leq f(\beta, N; I_2), \quad (\text{A1} \cdot 4)$$

$$V(I_1)g(\beta, \mu; I_1) \leq V(I')g(\beta, \mu; I') \leq V(I_2)g(\beta, \mu; I_2). \quad (\text{A1} \cdot 5)$$

Denoting as $\tilde{U}_N(x_1 \dots x_N)$ that which remains in the potential energy expression of the total system when we discard the terms associated with pairs of particles occupying more than one cube, we put

$$\tilde{\Phi}(\beta, N; I') \equiv \exp[N\tilde{f}(\beta, N; I')] = \frac{1}{N!} \int_{I'} \dots \int \exp(-\beta \tilde{U}_N) dx_1 \dots dx_N, \quad (\text{A1} \cdot 6)$$

$$\tilde{\Xi}(\beta, \mu; I') \equiv \exp[V(I')\tilde{g}(\beta, \mu; I')] = \sum_{n=1}^{\infty} e^{-\mu n} \exp[n\tilde{f}(\beta, n; I')], \quad (\text{A1} \cdot 7)$$

where I' denotes any of the three domains I_1 , I_2 and I' . These are the canonical and grandcanonical partition functions of three π -systems.

Van Hove already showed that

$$\varepsilon\left(1 - \frac{d_1}{d}\right) \leq f(\beta, N; I'') - \tilde{f}(\beta, N; I') \leq \frac{9}{2\pi} \beta \bar{u} \frac{V(I')}{N} \frac{d_1}{d_0^3 d}, \quad (\text{A1} \cdot 8)$$

where $\varepsilon(x)$ is a function with the property that

$$\lim_{x \rightarrow 1} \varepsilon(x) = 0. \quad (\text{A1} \cdot 9)$$

The expression of $\varepsilon(x)$ depends only on the nature of the interparticle force.

Therefore, we have

$$\exp[nf(\beta, n; I'')] = \exp[n\tilde{f}(\beta, n; I')] \exp(n\delta_n), \quad (\text{A1} \cdot 10)$$

$$|\delta_n| \leq \text{Max} \left[\varepsilon\left(1 - \frac{d_1}{d}\right), \frac{9}{2\pi} \beta \bar{u} \frac{V(I')}{n} \frac{d_1}{dd_0^3} \right], \quad (\text{A1} \cdot 11)$$

and thus we put

$$\mathcal{E} = \tilde{\mathcal{E}} \exp[V(\Gamma') \delta], \quad (\text{A1} \cdot 12)$$

where

$$|\delta| \leq \text{Max} \left[\frac{1}{v_0} \varepsilon \left(1 - \frac{d_1}{d} \right), \frac{9}{2\pi} \beta \bar{u} \frac{d_1}{dd_0^3} \right]. \quad (\text{A1} \cdot 13)$$

Thus we obtain

(L·1·1) :

$$|g(\beta, \mu; \Gamma') - \tilde{g}(\beta, \mu; \Gamma')| \leq \text{Max} \left[\frac{1}{v_0} \varepsilon \left(1 - \frac{d_1}{d} \right), \frac{9}{2\pi} \beta \bar{u} \frac{d_1}{dd_0^3} \right], \quad (\text{A1} \cdot 14)$$

$$|g(\beta, \mu; \Gamma') - \tilde{g}(\beta, \mu; \Gamma')| \leq \text{Max} \left[\frac{1}{v_0} \varepsilon \left(1 - \frac{d_1}{d} \right), \frac{9}{2\pi} \beta \bar{u} \frac{d_1}{dd_0^3} \right]. \quad (\text{A1} \cdot 15)$$

Supposing that Γ' is composed of cubes $\gamma_1, \gamma_2, \dots, \gamma_M$, we can write

$$\begin{aligned} \exp[V(\Gamma') \tilde{g}(\beta, \mu; \Gamma')] &= \{\exp[V(\gamma) \tilde{g}(\beta, \mu; \gamma)]\}^M \\ &= \exp[V(\Gamma') \tilde{g}(\beta, \mu; \gamma)], \end{aligned} \quad (\text{A1} \cdot 16)$$

showing that

$$\tilde{g}(\beta, \mu; \Gamma') = \tilde{g}(\beta, \mu; \gamma) \equiv \tilde{g}_a(\beta, \mu). \quad (\text{A1} \cdot 17)$$

Introducing this into the left-hand side of (L·1·1) and noting that $g(\beta, \mu; \Gamma')$ is independent of d , and the right-hand side is independent of $V(\Gamma')$, we can conclude that :

(L·1·2) : There exists

$$\lim_{d \rightarrow \infty} \tilde{g}_a(\beta, \mu) = \tilde{g}(\beta, \mu), \quad (\text{A1} \cdot 18)$$

and

$$\lim_{\Gamma(\Gamma) \rightarrow \infty} g(\beta, \mu; \Gamma) = \tilde{g}(\beta, \mu). \quad (\text{A1} \cdot 19)$$

It is to be noted that these convergences are uniform so long as β is bounded, because the right-hand side of (L·1·1) is bounded when β is bounded.

From (L·1·1) and (L·1·2), we can at once derive the following lemmas.

(L·1·3) : If

$$\lim_{\Gamma(\Gamma) \rightarrow \infty} V(\Gamma')/V(\Gamma) = 1,$$

$$\lim_{\Gamma(\Gamma) \rightarrow \infty} g(\beta, \mu; \Gamma') = \lim_{\Gamma(\Gamma) \rightarrow \infty} g(\beta, \mu; \Gamma). \quad (\text{A1} \cdot 20)$$

(L·1·4) : Let $g^*(\beta, \mu; \Gamma)$ be the corresponding quantity to $g(\beta, \mu; \Gamma)$ when the field of external forces with finite magnitude exists. Let Γ'' be the sub-space of Γ where the field is not zero.

$$\text{If} \quad \lim_{\Gamma(\Gamma) \rightarrow \infty} V(\Gamma'')/V(\Gamma) = 0,$$

we have

$$\lim_{\Gamma(\Gamma) \rightarrow \infty} g^*(\beta, \mu; \Gamma) = \lim_{\Gamma(\Gamma) \rightarrow \infty} g(\beta, \mu; \Gamma). \quad (\text{A1} \cdot 21)$$

Appendix 2. The proofs of lemmas of Sec. 4.

(L·2·1) : This is apparent if we note that

$$\begin{aligned}\log \tilde{\Xi}(\beta, \mu; \Gamma) &= M \tilde{g}_a(\beta, \mu), \\ &= N \frac{1}{\rho} \frac{1}{d^3} \tilde{g}_a(\beta, \mu),\end{aligned}\quad (\text{A2} \cdot 1)$$

and that β and μ are determined by

$$-\partial \tilde{g}_a(\beta, \mu) / \partial \beta = \rho d^3 \varepsilon, \quad -\partial \tilde{g}_a(\beta, \mu) / \partial \mu = \rho d^3. \quad (\text{A2} \cdot 2)$$

(L·2·2) : The application of central limit theorem to π -system proves the propositions. The detailed accounts of the theorem have been given in Sec. 3.

From (3·9) and (3·15) we have

$$\log \tilde{\mathcal{Q}}^a(E, N; \Gamma) = \beta E + \mu N + \log \tilde{\Xi}^a(\beta, \mu; \Gamma) + o(M). \quad (\text{A2} \cdot 3)$$

On the other hand

$$\log [kJE \text{ Min}_{E' \leq E' \leq E + \Delta E} \tilde{\mathcal{Q}}^a(E', N; \Gamma)] \leq \tilde{S}_R^a(E, N; \Gamma) \leq \log [kJE \text{ Max}_{E' \leq E' \leq E + \Delta E} \tilde{\mathcal{Q}}^a(E', N; \Gamma)]. \quad (\text{A2} \cdot 4)$$

Noting that $\log \tilde{\Xi}^a(\beta, \mu; \Gamma) = d^3 M \tilde{g}_a(\beta, \mu)$, we have

$$\tilde{S}_R^a(E, N; \Gamma) = \tilde{S}_G^a(E, N; \Gamma) + o(M). \quad (\text{A2} \cdot 5)$$

Quite analogously we can show that

$$\tilde{S}_R^a(E, N; \Gamma) = \tilde{S}_G^a(E, N; \Gamma) + o(M). \quad (\text{A2} \cdot 6)$$

(L·2·3) : Denoting the intracellular potential energy of Γ by E_0 , the intercellular potential energy of Γ by E_1 ($E = E_0 + E_1$), and the phase volume corresponding to

$$x < E_0 < x + dx, \quad y < E_1 < y + dy \quad (\text{A2} \cdot 7)$$

by $\omega(x, y, N) dx dy$, we have

$$\mathcal{Q}(E, N; \Gamma) = \int_{-\infty}^{\infty} \omega(E - y, y, N) dy, \quad (\text{A2} \cdot 8)$$

$$\tilde{\mathcal{Q}}^a(E, N; \Gamma) = \int_{-\infty}^{\infty} \omega(E, y, N) dy. \quad (\text{A2} \cdot 9)$$

Due to the nature of the interaction, we see that

$$\omega(x, y, N) = 0 \quad \text{when} \quad x < -cN\tilde{u}, \quad (\text{A2} \cdot 10)$$

$$\omega(x, y, N) = 0 \quad \text{when} \quad \infty > |y| > d(N, d) > 0, \quad (\text{A2} \cdot 11)$$

where

$$0 < d(N, d) / N = O(1/d). \quad (\text{A2} \cdot 12)$$

Thus

$$\begin{aligned} \int_E^{E+\Delta E} dE' \Omega(E', N; I') &\leq \int_{-\infty}^{E+\Delta E+\Delta(N, d)} dE' \int_{-\infty}^{\infty} \omega(E', \gamma, N) d\gamma \\ &= \int_{-\infty}^{E+\Delta E+\Delta(N, d)} dE' \tilde{\Omega}^d(E', N; I'). \end{aligned} \quad (\text{A2} \cdot 13)$$

Noting that

$$\tilde{\Omega}^d(E', N; I') = e^{(1/k) N \tilde{s}_c^d(E'/N, \rho) + o(M)} \quad (\text{A2} \cdot 14)$$

from (L·2·1) and (L·2·2),

$$\text{and that} \quad \partial \tilde{s}_c^d(\varepsilon, \rho) / \partial \varepsilon > 0 \quad \text{because} \quad \beta > 0, \quad (\text{A2} \cdot 15)$$

we have

$$\begin{aligned} k \log \left\{ \int_E^{E+\Delta E} dE' \Omega(E', N; I') \right\} &\leq N \tilde{s}_c^d([E + \Delta E + \Delta(N, d)]/N, \rho) + o(M) \\ &= N \{ \tilde{s}_c^d(E/N, \rho) + O(1/d) \} + o(M), \end{aligned} \quad (\text{A2} \cdot 16)$$

while

$$\int_E^{E+\Delta E} dE' \Omega(E', N; I') \geq \int_E^{E+\Delta E} dE' \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} dx \omega(E' - x, x, N), \quad (\text{A2} \cdot 17)$$

where

$$dE' \lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} dx \omega(E' - x, x) \quad (\text{A2} \cdot 18)$$

is the phase volume corresponding to the configuration where there is no intercellular interaction, and thus

$$\lim_{\varepsilon \rightarrow 0} \int_{-\varepsilon}^{\varepsilon} dx \omega(E' - x, x, N) \geq \tilde{\Omega}^d(E', N; I'). \quad (\text{A2} \cdot 19)$$

The definition of I' is the same as in § 3. From (L·1·2),

$$k \log \tilde{\Omega}^d(E', N; I') = N \{ \tilde{s}_c^d(E'/N, N/V) + O(1/d) \} + o(M). \quad (\text{A2} \cdot 20)$$

$$\therefore k \log \left\{ \int_E^{E+\Delta E} dE' \Omega(E', N; I') \right\} \geq N \{ \tilde{s}_c^d(\varepsilon, \rho) + O(1/d) \} + o(M). \quad (\text{A2} \cdot 21)$$

Again using (L·2·2), we obtain (L·2·3).

(L·2·4): The proof is directly obtained from Van Hove's paper and (L·1·1). We do not repeat them here.

Appendix 3.

We shall prove here the following theorem: If two real and finite numbers A and $L > 0$ are given, the equations

$$\begin{aligned}\partial \log \Xi / \partial \alpha + A &= 0 \\ \partial \log \Xi / \partial \lambda + L &= 0\end{aligned}\quad (\text{A3} \cdot 1)$$

have always a single pair of roots

$$\alpha = \theta > 0, \quad \lambda = \mu. \quad (\text{A3} \cdot 2)$$

Proof: Introducing a new function Ξ^* by

$$\Xi^*(A, L; \alpha, \lambda) = e^{\alpha A + \lambda L} \Xi(\alpha, \lambda), \quad (\text{A3} \cdot 3)$$

one can rewrite (A3.1) in the form:

$$\partial \log \Xi^* / \partial \alpha = 0, \quad \partial \log \Xi^* / \partial \lambda = 0. \quad (\text{A3} \cdot 4)$$

Thus the problem is to show that Ξ^* has a single minimum in the half-plane $\alpha > 0$, $-\infty < \lambda < +\infty$ which will be called σ -plane.

Now we can prove that Ξ^* tends to $+\infty$ at the whole boundary of σ .

(1) Boundary $\alpha = 0$, $\lambda \geq 0$.

Choosing a certain small integer $\delta > 0$, we have at once

$$\begin{aligned}\Xi^* &\geq e^{\lambda L} \sum_{N < L - \delta} \int_{-\infty}^{+\infty} e^{-\lambda N} \Omega(E, N) dE / N! h^{3N} \\ &\geq e^{\lambda \delta} \sum_{N < L - \delta} \int_{-\infty}^{+\infty} \Omega(E, N) dE / N! h^{3N} \rightarrow +\infty,\end{aligned}\quad (\text{A3} \cdot 5)$$

because the integration with respect to E extends to $+\infty$.

(2) Boundary $\alpha = 0$, $\lambda < 0$.

$$\begin{aligned}\Xi^* &\geq e^{\lambda L} \sum_{N > L + \delta} \int_{-\infty}^{+\infty} e^{-\lambda N} \Omega(E, N) dE / N! h^{3N} \\ &\geq e^{-\lambda \delta} \sum_{N > L + \delta} \int_{-\infty}^{+\infty} \Omega(E, N) dE / N! h^{3N} \rightarrow +\infty.\end{aligned}\quad (\text{A3} \cdot 6)$$

(3) Boundary $\lambda = +\infty$, α is arbitrary (including $+\infty$).

$$\begin{aligned}\Xi^* &\geq e^{\alpha A + \lambda L} \sum_{N < L - \delta} \int_{-\infty}^A e^{-\alpha E - \lambda N} \Omega(E, N) dE / N! h^{3N} \\ &\geq e^{\lambda \delta} \sum_{N < L - \delta} \int_{-\infty}^A \Omega(E, N) dE / N! h^{3N} \xrightarrow{(\lambda \rightarrow +\infty)} +\infty.\end{aligned}\quad (\text{A3} \cdot 7)$$

(4) Boundary $\lambda = -\infty$, α is arbitrary (including $+\infty$).

$$\Xi^* \geq e^{-\lambda \delta} \sum_{N > L + \delta} \int_{-\infty}^A \Omega(E, N) dE / N! h^{3N} \xrightarrow{(\lambda \rightarrow -\infty)} +\infty. \quad (\text{A3} \cdot 8)$$

(5) Boundary $\alpha = +\infty$, λ is finite.

Choosing a certain small energy $\varepsilon > 0$, one obtains

$$\begin{aligned}\Xi^* &> e^{\alpha A + \lambda L} \sum_N \int_{-\infty}^{A-\varepsilon} e^{-\alpha E - \lambda N} Q(E, N) dE / N! h^{3N} \\ &> e^{\alpha A + \lambda L} \sum_N \int_{-\infty}^{A-\varepsilon} e^{-\lambda N} Q(E, N) dE / N! h^{3N} \xrightarrow{(\alpha \rightarrow +\infty)} + \infty.\end{aligned}\quad (\text{A3}\cdot 9)$$

In view of (A3·5) — (A3·9), it will be obvious that our theorem holds when and only when the second derivative of $\log \Xi^*$ is always positive. To show this, let us differentiate $\log \Xi^*$ twice in the direction s which makes an angle ϕ with λ axis at an arbitrary point on σ . Then we have

$$\partial^2 \log \Xi^* / \partial s^2 = \partial^2 \log \Xi^* / \partial \alpha^2 \cdot \sin^2 \phi + 2 \partial^2 \log \Xi^* / \partial \alpha \partial \lambda \cdot \sin \phi \cos \phi + \partial^2 \log \Xi^* / \partial \lambda^2 \cdot \cos^2 \phi. \quad (\text{A3}\cdot 10)$$

Remembering (2·4) and (2·5) in the text, the required relation can at once be derived from (A3·9):

$$\begin{aligned}\frac{\partial^2 \log \Xi^*}{\partial s^2} &= \frac{1}{\Xi(\alpha, \lambda)} \sum_N \left[\left(E + \frac{\partial \log \Xi}{\partial \alpha} \right) \sin \phi + \left(N + \frac{\partial \log \Xi}{\partial \lambda} \right) \cos \phi \right]^2 \times \\ &\quad \times e^{-\alpha E - \lambda N} Q(N, E) \frac{dE}{N! h^{3N}} > 0.\end{aligned}\quad (\text{A3}\cdot 11)$$

References

- 1) Van Hove, *Physica* **15** (1949), 951.
- 2) S. Katsura, *Prog. Theor. Phys.* **11** (1954), 476; *Proc. Int. Conf. Theor. Phys. Kyoto*, 534, (1953).
- 3) S. Katsura, *Prog. Theor. Phys.* **13** (1955), 571.
- 4) I. Khinchin (translated by Gamow), *Mathematical Foundations of Statistical Mechanics*, New York, 1949.
- 5) L. Onsager and S. Machlup, *Phys. Rev.* **91** (1953), 1505, 1512.
- 6) N. Hashitsume, *Prog. Theor. Phys.* **8** (1952), 461; *Proc. Int. Conf. Theor. Phys. Kyoto*, 495, (1953).
- 7) M. S. Green, *J. Chem. Phys.* **20** (1952), 1281.
- 8) T. Yamamoto, *Prog. Theor. Phys.* **10** (1953), 11.

A Relativistic Wave Equation for a Particle with Two Mass States of Spin 1 and 0

Haruo SHIMAZU

Department of Physics, Yokohama National University, Yokohama

(Received April 4, 1956)

A relativistic wave equation is given which describes a particle having two different mass states of spin 1 and 0. The energy density is positive definite and the usual method of quantization is applicable.

§ 1. Introduction

Recently, Bhabha¹⁾ has given a relativistic wave equation which describes a particle with two mass states, one of spin $3/2$ and the other spin $1/2$. This equation was reformulated by Gupta²⁾ in the analogous way to Rarita-Schwinger's formula and the quantization was performed.

The equation of Bhabha is

$$(\alpha_\mu p^\mu - \chi) \psi(x) = 0, \quad (1.1)$$

where $p^\mu = -i\partial/\partial x_\mu$, $x_\mu = (x_0, x_1, x_2, x_3)$, $x_0 \equiv ct$, and χ is a constant related to masses. α_μ are the 20×20 matrices. Raising and lowering of tensor indices are done by the metric tensor $g_{\mu\nu} = g^{\mu\nu}$ defined by

$$g_{00} = -g_{11} = -g_{22} = -g_{33} = 1, \quad g_{\mu\nu} = 0 \text{ for } \mu \neq \nu.$$

Before this equation was found, the equation which have the form

$$(\alpha_\mu p^\mu - \chi) \psi(x) = 0, \quad (1.2)$$

where α_μ are certain square matrices, had been investigated by various authors³⁾⁻⁷⁾, and it was shown that these equations have solutions describing particles with various mass and spin states, neither the energy nor the charge being, however, made to be positive definite. As is well known, the usual method of quantization is physically significant only when the energy or the charge is definite. But the equations treated by these authors were not sufficiently general and Bhabha has shown that there really exists an equation which describes a Fermi particle with two mass states and definite charge density.

His theory would be of interest at the present time, because a number of new particles have been observed recently, and some of them of half-integral spin may be described by such multi-mass equations. The situation is similar for particles of integral spin. We shall, therefore, extend the theory to Bose particles, giving an equation which describes

a particle having two different mass states of spin 1 and 0. The equation is defined in § 2 and its properties are investigated in § 3, showing that the energy density is positive definite. § 4 is devoted to the quantization of the theory.

§ 2. The wave equation

At first we give a brief description about the properties of the equation having the form (1.2), as far as it is necessary for the development of this paper. The equation (1.2) is relativistically invariant if, with the Lorentz transformation

$$x'_\mu = u_\mu^\nu x_\nu, \quad p'_\mu = u_\mu^\nu p_\nu, \quad (2.1)$$

there exists a transformation

$$\psi'(x') = U\psi(x), \quad (2.2)$$

such that

$$\alpha'_\mu = u_\mu^\nu \alpha_\nu = U^{-1} \alpha_\mu U. \quad (2.3)$$

The Lorentz rotation can be built up from the infinitesimal transformations

$$x'_\mu = x_\mu + \varepsilon_{\mu\nu} x^\nu, \quad U = 1 + 1/2 \varepsilon^{\rho\sigma} I_{\rho\sigma}, \quad (2.4)$$

where

$$\varepsilon_{\mu\nu} + \varepsilon_{\nu\mu} = 0, \quad I_{\rho\sigma} + I_{\sigma\rho} = 0.$$

The condition (2.3), applied to (2.4), gives

$$[\alpha_\mu, I_{\rho\sigma}] = \delta_{\mu\rho} \alpha_\sigma - \delta_{\mu\sigma} \alpha_\rho. \quad (2.5)$$

If one puts $U=R$ for the reflection of the three spatial axes, (2.3) requires ($k=1, 2, 3$)

$$R^{-1} \alpha_0 R = \alpha_0, \quad R^{-1} \alpha_k R = -\alpha_k. \quad (2.6)$$

$I_{\mu\nu}$ and R also satisfy the following commutation rules ($k, l=1, 2, 3$),

$$[I_{\mu\nu}, I_{\rho\sigma}] = -\delta_{\mu\rho} I_{\nu\sigma} + \delta_{\nu\sigma} I_{\mu\rho} + \delta_{\mu\sigma} I_{\nu\rho} - \delta_{\nu\rho} I_{\mu\sigma},$$

$$R^{-1} I_{0k} R = -I_{0k}, \quad R^{-1} I_{kl} R = I_{kl}. \quad (2.7)$$

The wave equation (1.2) may be derived from the Lagrangian density

$$L(x) = \psi^*(x) D(\alpha_\mu p^\mu - \chi) \psi(x), \quad (2.8)$$

provided there exists a matrix D , which is non-singular and such that $\alpha_\mu^* D = D \alpha_\mu$. The invariance of (1.2) requires

$$U^* D U = D. \quad (2.9)$$

By use of (2.4) for U , one gets from (2.9)

$$I_{\rho\sigma}^* D + D I_{\rho\sigma} = 0, \quad (2.10)$$

and

$$R^*D=DR. \quad (2 \cdot 11)$$

The charge-current four vector and the energy-momentum tensor are defined by

$$\begin{aligned} s_\mu &= \psi^* D\alpha_\mu \psi, \\ t_{\mu\nu} &= 1/2[\psi^* D\alpha_\mu p_\nu \psi - (p_\nu \psi^*) D\alpha_\mu \psi]. \end{aligned} \quad (2 \cdot 12)$$

All the formulations are equivalent under the transformation by an arbitrary non-singular matrix s ,

$$\begin{aligned} \psi &\rightarrow s^{-1} \psi, \\ D &\rightarrow s^* D s, \\ \alpha_\mu &\rightarrow s^{-1} \alpha_\mu s, \\ U &\rightarrow s^{-1} U s. \end{aligned} \quad (2 \cdot 13)$$

To construct the representation of α_μ , we shall work with the two sets (K_1, K_2, K_3) and (L_1, L_2, L_3) defined by

$$\begin{aligned} K_1 &= \frac{1}{2}(i I_{23} + I_{01}), & K_2 &= \frac{1}{2}(i I_{31} + I_{02}), & K_3 &= \frac{1}{2}(i I_{12} + I_{03}), \\ L_1 &= \frac{1}{2}(i I_{23} - I_{01}), & L_2 &= \frac{1}{2}(i I_{31} - I_{02}), & L_3 &= \frac{1}{2}(i I_{12} - I_{03}), \end{aligned} \quad (2 \cdot 14)$$

both of which transform as components of a vector under spatial rotations, and obey the commutation rules

$$[K_1, K_2] = iK_3, \quad [L_1, L_2] = iL_3, \quad (2 \cdot 15)$$

the remaining being obtained by cyclic permutations of $(1, 2, 3)$. Then $K^2 = K_1^2 + K_2^2 + K_3^2$ and $L^2 = L_1^2 + L_2^2 + L_3^2$ commute with all the $I_{\mu\nu}$ and their eigenvalues are $k(k+1)$ and $l(l+1)$, respectively, where k and l are positive numbers, integral or half-integral, or zero. $S_1 = iI_{23}$, $S_2 = iI_{31}$ and $S_3 = iI_{12}$ also satisfy the similar relation to (2.15) and can be interpreted as components of the spin angular momentum, $S^2 = S_1^2 + S_2^2 + S_3^2$ having the eigenvalue $s(s+1)$. To investigate the relation between the mass and spin eigenvalues, it is convenient to use the representation in which S and S_z are diagonal as well as K and L . D can also be given in the diagonal form in this representation. Corresponding to the given k and l , the values of s form a progression

$$s_{max} (=k+l), s_{max}-1, \dots, s_{min} (=|k-l|),$$

and D is of the form

$$(s|D|s') = d(-1)^s \delta_{ss'},$$

where d is a constant. One may assume $d = \pm 1$ without loss of generality, the sign being given arbitrarily for any particular k and l values.

The α_μ are determined uniquely when one defines the transformation properties of ψ , i.e., $I_{\mu\nu}$ and R , and fixes the arbitrariness of the sign in the representation of D .

As for the equation with two states of spin 1 and 0, we shall consider the follow-

ing transformation properties, which would be the simplest for our purpose, even if another choice may not be excluded.

$$I_{\mu\nu} = \begin{pmatrix} I_{\mu\nu}(1, 1) & \cdot & \cdot & \cdot & \cdot \\ \cdot & I_{\mu\nu}(1/2, 1/2) & \cdot & \cdot & \cdot \\ \cdot & \cdot & I_{\mu\nu}(0, 0) & \cdot & \cdot \\ \cdot & \cdot & \cdot & I_{\mu\nu}(1/2, 1/2) & \cdot \\ \cdot & \cdot & \cdot & \cdot & I_{\mu\nu}(0, 0) \end{pmatrix}, \quad (2 \cdot 17)$$

where $I_{\mu\nu}(k, l)$ are irreducible elements of $I_{\mu\nu}$ specified by k and l , and

$$R = \begin{pmatrix} A_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & A_1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & A_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & A_1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & A_0 \end{pmatrix}, \quad (2 \cdot 18)$$

where

$$A_2 = \begin{pmatrix} \mathbf{I}_5 & \cdot & \cdot \\ \cdot & -\mathbf{I}_3 & \cdot \\ \cdot & \cdot & \mathbf{I} \end{pmatrix}, \quad A_1 = \begin{pmatrix} -\mathbf{I}_3 & \cdot \\ \cdot & \mathbf{I} \end{pmatrix}, \quad A_0 = \mathbf{I}.$$

\mathbf{I}_n stands for the $n \times n$ unit matrix. Corresponding to the $I_{\mu\nu}$ and R given above, D can be given in the diagonal form as follows.

$$D = \begin{pmatrix} d_1 A_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & d_2 A_1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & d_3 A_0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & d_4 A_1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & d_5 A_0 \end{pmatrix}, \quad (2 \cdot 19)$$

where d_m ($m=1, 2, 3, 4, 5$) take the value $+1$ or -1 independently of each other. We choose them as follows.

$$d_1 = d_2 = -1, \quad d_3 = d_4 = d_5 = 1. \quad (2 \cdot 20)$$

α_μ can now be determined uniquely from the above definitions. α_0 is given by the relations

$$[\alpha_0, I_{kl}] = 0, \quad [[\alpha_0, I_{0k}], I_{0k}] = \alpha_0, \quad (2 \cdot 21)$$

α_k being defined by

$$[\alpha_0, I_{0k}] = \alpha_k. \quad (2 \cdot 22)$$

For the representation $I_{\mu\nu}(k, l)$ and the method of obtaining the representation of α_μ by use of (2.21) and (2.22), one may refer to Reference (7).

The α_μ matrices are thus given as follows. (The symbol* denotes the Hermitian conjugate of a matrix.)

$$\alpha_0 = \begin{pmatrix} \cdot & u_0 & \cdot & fu_0 & \cdot \\ u_0^* & \cdot & av_0 & \cdot & dv_0 \\ \cdot & -av_0^* & \cdot & cv_0^* & \cdot \\ -fu_0^* & \cdot & cv_0 & \cdot & bv_0 \\ \cdot & -dv_0^* & \cdot & bv_0^* & \cdot \end{pmatrix}, \quad \alpha_k = \begin{pmatrix} \cdot & u_k & \cdot & fu_k & \cdot \\ -u_k^* & \cdot & av_k & \cdot & dv_k \\ \cdot & av_k^* & \cdot & -cv_k^* & \cdot \\ fu_k^* & \cdot & cv_k & \cdot & bv_k \\ \cdot & dv_k^* & \cdot & -bv_k^* & \cdot \end{pmatrix},$$

$$u_0 = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \sqrt{3}/2 \end{pmatrix}, \quad v_0 = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix},$$
$$u_1 = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{6} & \cdot \\ \cdot & -1/\sqrt{2} & \cdot & \cdot \\ \cdot & \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot & 1/\sqrt{2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1/\sqrt{2} \\ \sqrt{3}/6 & \cdot & -\sqrt{3}/6 & \cdot \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ \cdot \\ \cdot \end{pmatrix}, \tag{2.23}$$

$u_2 = iu_1, \quad v_2 = -iv_1,$

$$u_3 = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot \\ \cdot & 2/\sqrt{3} & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \\ \cdot & 1/\sqrt{6} & \cdot & \cdot \end{pmatrix}, \quad v_3 = \begin{pmatrix} \cdot \\ -1 \\ \cdot \\ \cdot \end{pmatrix},$$

the rows and columns being distinguished by s and $m(=s_z)$ as well as k and l . They are arranged in the order such that the larger s appears to the left of the smaller, and once s is given the larger m appears to the left of the smaller. The first nine rows and columns are thus arranged in the order, denoting them by (s, m) for the particular k and l values

$(2, 2) \ (2, 1), \ (2, 0), \ (2, -1), \ (2, -2), \ (1, 1), \ (1, 0), \ (1, -1), \ (0, 0),$

and similarly for the others.

For obtaining the spin values and the corresponding rest masses, one writes the matrix elements of α_0 for each spin state separately,

$\lambda_2 \equiv (s=2 | \alpha_0 | s=2) = 0, \tag{2.24}$

$\lambda_1 \equiv (s=1 | \alpha_0 | s=1) = \begin{pmatrix} \cdot & 1 & f \\ 1 & \cdot & \cdot \\ -f & \cdot & \cdot \end{pmatrix}, \tag{2.25}$

$$\lambda_0 \equiv (s=0|\alpha_0|s=0) = \begin{pmatrix} \cdot & \sqrt{3/2} & \cdot & \sqrt{3/2}f & \cdot \\ \sqrt{3/2} & \cdot & a & \cdot & d \\ \cdot & -a & \cdot & c & \cdot \\ -\sqrt{3/2}f & \cdot & c & \cdot & b \\ \cdot & -d & \cdot & b & \cdot \end{pmatrix}. \quad (2 \cdot 26)$$

Here a, b, c, d and f are numbers which can not be determined from the transformation properties. Strictly speaking, each matrix element for $s > 0$ in the above expression should be written as a diagonal matrix of $2s+1$ rows and columns with same diagonal elements. But we have simply put them as numbers, which has nothing to do with minimal equations. (2·24) shows that the state of spin 2 is not realized. One can make the energy density positive definite by putting certain restrictions on the constants.

We shall define the constants by

$$\begin{aligned} c &= d = 0, \\ 3f^2 a^2 + 3b^2 &= 2a^2 b^2, \\ 3a^2 / (2a^2 - 3) > b^2 > \frac{2}{3} a^2 (2a^2 - 3) > 0, \end{aligned} \quad (2 \cdot 27)$$

which make the minimal equations for matrices λ_1 and λ_0 be

$$\begin{aligned} \lambda_1 [\lambda_1^2 - (1 - f^2)] &= 0, \\ \lambda_0 [\lambda_0^2 - (\frac{3}{2} - a^2 + b^2 - \frac{3}{2} f^2)] &= 0, \end{aligned} \quad (2 \cdot 28)$$

respectively. (2·28) shows that there are two states of spin 1 and 0, and their masses are, respectively,

$$m_1 = \frac{\chi}{(1 - f^2)^{1/2}}, \quad m_0 = \frac{\chi}{(\frac{3}{2} - a^2 + b^2 - \frac{3}{2} f^2)^{1/2}}. \quad (2 \cdot 29)$$

Then the constants must satisfy the inequalities,

$$f^2 < 1, \quad \frac{3}{2} - a^2 + b^2 - \frac{3}{2} f^2 > 0,$$

which are satisfied by (2·27).

§ 3. Properties of the equation

We shall now prove that the energy density for the equation is positive definite. The equation contains the constants not determined from the transformation properties, and we have defined them by (2·27). Though we can put more restrictions besides them, it will become clear that they are sufficient for making the energy density positive definite.

To see this, we first transform α_0 to the pseudo-diagonal form by the transformation (2·13) with the transformation function s which is given below.

For the state of spin 1, one can transform λ_1 such that

$$s^{-1} \lambda_1 s = \begin{pmatrix} \hat{\xi} & \cdot & \cdot \\ \cdot & -\hat{\xi} & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}, \quad (3.1)$$

where $\hat{\xi} = (1 - f^2)^{1/2}$. From this it follows that but for an arbitrary multiplying factor

$$s_{11} = 1, \quad s_{21} = 1/\hat{\xi}, \quad s_{31} = -f/\hat{\xi}, \\ s_{k2}(\hat{\xi}) = s_{k1}(-\hat{\xi}). \quad (3.2)$$

Then D becomes after this transformation

$$s^*(s=1|D|s=1)s = s^* \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & -1 \end{pmatrix} s = \begin{pmatrix} \alpha & \cdot & \cdot \\ \cdot & \alpha & \cdot \\ \cdot & \cdot & \delta \end{pmatrix}, \quad (3.3)$$

where

$$\alpha = |s_{11}|^2 + |s_{21}|^2 - |s_{31}|^2, \quad (3.4)$$

and the value of δ is not necessary. Substituting (3.2) in (3.4) one gets $\alpha = 2$. In this representation the solution of $(\alpha_0 p^0 - \chi)\psi(x) = 0$, the equation corresponding to a particle at rest, is without normalization

$$\psi = p\phi_+ e^{i\chi x_0/\xi} + q\phi_- e^{-i\chi x_0/\xi}, \quad (3.5)$$

where

$$\phi_+ = \begin{pmatrix} 1 \\ \cdot \\ \cdot \end{pmatrix}, \quad \phi_- = \begin{pmatrix} \cdot \\ 1 \\ \cdot \end{pmatrix}, \quad (3.6)$$

and p and q are arbitrary constants. By use of (2.12), (3.1), (3.3), (3.5) and (3.6) the energy density reduces to

$$t_{00} = \alpha\chi(|p|^2 + |q|^2), \quad (3.7)$$

showing that the sign of the energy density is equal to that of α (assuming $\chi > 0$), and positive in this case.

The similar procedure is adopted for the state of spin 0. The transformation function s , which corresponds (3.1) for the state of spin 1, is determined to satisfy

$$s^{-1} \lambda_0 s = \begin{pmatrix} \eta & \cdot & \cdot \\ \cdot & -\eta & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}, \quad (3.8)$$

where

$$\eta = (\frac{3}{2} - a^2 + b^2 - \frac{3}{2}f^2)^{1/2}.$$

From this it follows that but for an arbitrary multiplying factor

$$s_{11} = \sqrt{2/3} (1/\eta) (\eta^2 + a^2), \quad s_{21} = 1, \quad s_{31} = -a/\eta,$$

$$s_{41} = \frac{1}{f} \left[\frac{2}{3} (\eta^2 + a^2) - 1 \right], \quad s_{51} = \frac{b}{f\eta} \left[\frac{2}{3} (\eta^2 + a^2) - 1 \right], \quad (3.9)$$

$$s_{k2}(\eta) = s_{k1}(-\eta).$$

D becomes after this transformation

$$s^*(s=0|D|s=0)s = s^* \begin{pmatrix} -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & -1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix} s = \begin{pmatrix} \beta & \cdot & \cdot & \cdot & \cdot \\ \cdot & \beta & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} s, \quad (3.10)$$

where

$$\beta = -|s_{11}|^2 - |s_{21}|^2 + |s_{31}|^2 + |s_{41}|^2 + |s_{51}|^2, \quad (3.11)$$

and \hat{o} is a 3×3 matrix, the form of which is not necessary for our purpose. The solution corresponding to a particle at rest is

$$\phi = p\phi_+ e^{i\chi x_0/\eta} + q\phi_- e^{-i\chi x_0/\eta}, \quad (3.12)$$

where

$$\phi_+ = \begin{pmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad \phi_- = \begin{pmatrix} \cdot \\ 1 \\ \cdot \\ \cdot \\ \cdot \end{pmatrix}, \quad (3.13)$$

and p and q are arbitrary constants. The energy density is

$$t_{00} = \beta \chi (|p|^2 + |q|^2). \quad (3.14)$$

In order that this is positive, it is necessary that $\beta > 0$. Substituting (3.9) in (3.11) and using (2.27), one gets

$$\beta = \frac{1}{\eta^2} \left[2a^2 - 3 + \frac{3b^2}{a^4} \left(-\frac{3b^2}{2a^2 - 3} - 2a^2 \right) \right],$$

which is positive from the definition (2.27).

Thus it becomes clear that the equation has the solution with two states, one of mass $\chi/\tilde{\eta}$ and spin 1, and the other of mass χ/η and spin 0, the energy density being positive definite.

It remains to be proved that the equation is irreducible, that is, the α_μ can not be decomposed into the direct sum of irreducible sets. This may be proved in the way: we assume that the α_μ are decomposable and show later that the assumption leads to a contradiction. The procedure is similar to that given by Bhabha when he gave the equation with two mass states for a Fermi particle, so we do not give the proof in details but a few remarks.

As the equation contains two states and the energy density is definite, the α_μ would have the form

$$\alpha_{\mu} = \begin{bmatrix} \alpha_{\mu}^{(1)} \\ \alpha_{\mu}^{(2)} \end{bmatrix},$$

if they were decomposable, in which $\alpha_{\mu}^{(1)}$ and $\alpha_{\mu}^{(2)}$ transform under Lorentz transformation according to the representations $I_{\mu\nu}(1, 1) + I_{\mu\nu}(\frac{1}{2}, \frac{1}{2}) + I_{\mu\nu}(0, 0)$ and $I_{\mu\nu}(\frac{1}{2}, \frac{1}{2}) + I_{\mu\nu}(0, 0)$, respectively. The latter is the same with that of Kemmer's equation for spin 0, but the former differs from Kemmer's for spin 1. Though the equation containing the representation $I_{\mu\nu}(1, 1) + I_{\mu\nu}(\frac{1}{2}, \frac{1}{2}) + I_{\mu\nu}(0, 0)$ has in general two solutions of spin 1 and 0 and the energy density is not definite, the special choice of the constants appearing in the equation which are independent of transformation properties would make the amplitude corresponding to the state of spin 0 vanish. Such a situation would become clear in the Appendix.

Our equation is, however, indecomposable in such a way.

§ 4. Quantization and concluding remarks

Our equation can be quantized in the similar way to that for the Bose particle with one mass state. The commutation relation between the components of the field for a given time is⁴⁾

$$[D\alpha_0 \psi(x_0, \mathbf{x}), \psi^*(x_0, \mathbf{x}') D\alpha_0] = D\alpha_0 \delta(\mathbf{x} - \mathbf{x}'), \quad (4.1)$$

independent of the representation.

ψ and ψ^* may be expanded as follows.

$$\begin{aligned} \psi(x_{\mu}) &= 1/\sqrt{V} \sum_{m, r, k} [U_+(m, r, \mathbf{k}) \phi_+(m, r, \mathbf{k}) e^{ik_{\mu}x^{\mu}} \\ &\quad + U_-^*(m, r, \mathbf{k}) \phi_-(m, r, \mathbf{k}) e^{-ik_{\mu}x^{\mu}}], \\ \psi^*(x_{\mu}) &= 1/\sqrt{V} \sum_{m, r, k} [U_+^*(m, r, \mathbf{k}) \phi_+^*(m, r, \mathbf{k}) e^{-ik_{\mu}x^{\mu}} \\ &\quad + U_-(m, r, \mathbf{k}) \phi_-^*(m, r, \mathbf{k}) e^{ik_{\mu}x^{\mu}}]. \end{aligned} \quad (4.2)$$

$\phi_+(m, r, \mathbf{k}) e^{ik_{\mu}x^{\mu}}$ and $\phi_-(m, r, \mathbf{k}) e^{-ik_{\mu}x^{\mu}}$ are the momentum eigenfunctions corresponding to the mass m and the parameter r used to specify the polarization, and reduce in the rest system to $\phi_+ e^{i\chi x_0/\eta}$ and $\phi_- e^{-i\chi x_0/\eta}$ in (3.6) or $\phi_+ e^{i\chi x_0/\eta}$ and $\phi_- e^{-i\chi x_0/\eta}$ in (3.13) but for a multiplying factor, according as $m=m_1$ or $m=m_0$.

They can be normalized to

$$\phi_{\varepsilon}^*(m, r, \mathbf{k}) D\alpha_0 \phi_{\varepsilon'}(m', r', \mathbf{k}) = \varepsilon \delta_{mm'} \delta_{rr'} \delta_{\varepsilon\varepsilon'}. \quad (4.3)$$

It is because that as ϕ_+ and ϕ_- in (3.6) and (3.13) satisfy (4.3) but for a numerical factor, (4.3) may hold in any frame of reference due to the relativistic invariance of it.

From (4.3) it follows that

$$\sum_{m, r} D\alpha_0 \phi_{\varepsilon}(m, r, \mathbf{k}) \phi_{\varepsilon}^*(m, r, \mathbf{k}) D\alpha_0 = \varepsilon D\alpha_0. \quad (4.4)$$

Combining (4.1) with (4.4) one gets

$$[U_{\varepsilon}(m, r, k), U_{\varepsilon'}^*(m', r', k')] = \delta_{mm'} \delta_{rr'} \delta_{kk'} \delta_{\varepsilon\varepsilon'}. \quad (4.5)$$

As $U_{\varepsilon}^*(m, r, k) U_{\varepsilon}(m, r, k)$ are positive integers or zero, the energy of the free field, given by

$$\begin{aligned} E &= \int_{t_{00}} d^3x = \frac{-i}{2} \int d^3x \left[\psi^* D\alpha_0 \partial_0 \psi - (\partial_0 \psi^*) D\alpha_0 \psi \right] \\ &= \sum_{m,r,k} k_0 [U_+^*(m, r, k) U_+(m, r, k) + U_-(m, r, k) U_-^*(m, r, k)], \quad (4.6) \end{aligned}$$

is positive definite.

The method of quantization given above may not be restricted to the two mass equation, but will be equally applicable to more general cases of multimass equations with various spin states.

The interaction with the electromagnetic field is introduced by replacing p_{μ} by $p_{\mu} - eA_{\mu}$, by which the transition between the two states occurs, the transition probability depending essentially on the choice of the constants which are left to be determined. The state of spin 1(0) is then considered as "the excited state" of that of spin 0(1).

Though whether such particles will prove to exist in nature is an open question, the equation for a particle with more than one mass states such as given in this paper may be considered from the formalistic view point as the natural extension of usual ones. The theory is, however, unsatisfactory in the sense that the constants appeared in the equation are not determined uniquely from the values of mass and spin, and no principle to define them can be given now.

Appendix

As the equation describes a Bose particle it may be written down in tensor form. This problem is treated here. According to (2.17) for the representation of $I_{\mu\nu}$, ψ is represented by

$$\psi = \begin{pmatrix} \psi(1,1) \\ \psi(1/2, 1/2) \\ \bar{\psi}(0,0) \\ \bar{\psi}(1/2, 1/2) \\ \bar{\psi}(0,0) \end{pmatrix}. \quad (A.1)$$

As $I_{\mu\nu}(1/2, 1/2)$ and $I_{\mu\nu}(0, 0)$ appear twice in (2.17), we have distinguished one by attaching the bar from the other in the components of ψ in (A.1). ψ has 19 components which are distinguished by s and m besides k and l corresponding to the representation of $I_{\mu\nu}$. We shall denote the components of $\psi(k, l)$ by $\psi_m^s(k+l)$.

One introduces a second rank symmetric tensor $\theta^{\mu\nu}$, two vectors $\varphi_{(1)}^{\mu}$ and $\varphi_{(2)}^{\mu}$, and two scalars $U_{(1)}$ and $U_{(2)}$, which are related to $\psi_m^s(k+l)$ by

$$\begin{aligned} \psi_2^2(2) - \psi_{-2}^2(2) &= -2i\theta^{12}, & \psi_2^2(2) + \psi_{-2}^2(2) &= \theta^{11} - \theta^{22}, \\ \psi_1^2(2) - \psi_{-1}^2(2) &= -2\theta^{13}, & \psi_1^2(2) + \psi_{-1}^2(2) &= 2i\theta^{23}, \end{aligned}$$

$$\begin{aligned}
 \psi_0^2(2) &= -(1/\sqrt{6})\theta^{11} - (1/\sqrt{6})\theta^{22} + \sqrt{2/3}\theta^{33}, \\
 \psi_1^1(2) - \psi_{-1}^1(2) &= 2\theta^{01}, \quad \psi_1^1(2) + \psi_{-1}^1(2) = -2i\theta^{02}, \\
 \psi_0^1(2) &= -\sqrt{2}\theta^{03}, \quad \psi_0^0(2) = \sqrt{3/2}(\theta^{00} + \frac{1}{3}\theta^{11} + \frac{1}{3}\theta^{22} + \frac{1}{3}\theta^{33}), \\
 \psi_1^1(1) - \psi_{-1}^1(1) &= \varphi_{(1)}^1, \quad \psi_1^1(1) + \psi_{-1}^1(1) = -i\varphi_{(1)}^2, \\
 \psi_0^1(1) &= -(1/\sqrt{2})\varphi_{(1)}^3, \quad \psi_0^0(1) = (1/\sqrt{2})\varphi_{(1)}^0, \quad \psi_0^0(0) = U_{(1)}, \\
 \bar{\psi}_1^1(1) - \bar{\psi}_{-1}^1(1) &= \varphi_{(2)}^1, \quad \bar{\psi}_1^1(1) + \bar{\psi}_{-1}^1(1) = -i\varphi_{(2)}^2, \\
 \bar{\psi}_0^1(1) &= -(1/\sqrt{2})\varphi_{(2)}^3, \quad \bar{\psi}_0^0(1) = (1/\sqrt{2})\varphi_{(2)}^0, \quad \bar{\psi}_0^0(0) = U_{(2)}.
 \end{aligned} \tag{A.2}$$

As for the new variables it follows that

$$\begin{aligned}
 \chi^{\theta\mu\nu} &= \frac{1}{2}[p^\mu \varphi_{(1)}^\nu + p^\nu \varphi_{(1)}^\mu + f(p^\mu \varphi_{(2)}^\nu + p^\nu \varphi_{(2)}^\mu)], \\
 \chi \varphi_{(1)}^\mu &= 2p_\nu \theta^{\mu\nu} + \sqrt{2} a p^\mu U_{(1)} - \frac{1}{2} p^\mu \theta, \\
 \chi \varphi_{(2)}^\mu &= -2f p_\nu \theta^{\mu\nu} + \sqrt{2} b p^\mu U_{(2)} + (f/2) p^\mu \theta, \\
 \chi U_{(1)} &= -(a/\sqrt{2}) p_\mu \varphi_{(1)}^\mu, \\
 \chi U_{(2)} &= (b/\sqrt{2}) p_\mu \varphi_{(2)}^\mu,
 \end{aligned} \tag{A.3}$$

where $\theta = \sum \theta_{\mu}^{\mu}$. From these equations, it is easy to show that ψ satisfies the equation

$$(\square - m_1^2)(\square - m_2^2)\psi = 0. \tag{A.4}$$

It is to be noted that the quantization is not easily applicable to (A.3), but only to the original equation.

If one puts $f=0$, the equation is not irreducible, but is decomposed into two irreducible sets, one consists of $\theta^{\mu\nu}$, $\varphi_{(1)}^\mu$ and $U_{(1)}$, and the other $\varphi_{(2)}^\mu$ and $U_{(2)}$. In this case it follows that in one set

$$\begin{aligned}
 \chi \theta^{\mu\nu} &= \frac{1}{2}(p^\mu \varphi_{(1)}^\nu + p^\nu \varphi_{(1)}^\mu), \\
 \chi \varphi_{(1)}^\mu &= 2p_\nu \theta^{\mu\nu} + \sqrt{2} a p^\mu U_{(1)} - \frac{1}{2} p^\mu \theta, \\
 \chi U_{(1)} &= -(a/\sqrt{2}) p_\mu \varphi_{(1)}^\mu,
 \end{aligned} \tag{A.5}$$

and in the other

$$\begin{aligned}
 \chi \varphi_{(2)}^\mu &= \sqrt{2} b p^\mu U_{(2)}, \\
 \chi U_{(2)} &= (b/\sqrt{2}) p_\mu \varphi_{(2)}^\mu.
 \end{aligned} \tag{A.6}$$

(A.6) is identical with the scalar equation given by Kemmer. To investigate the properties of the equations (A.5), one multiplies the second with p_μ and obtains

$$p_\mu \varphi_{(1)}^\mu = \sqrt{2}/a \chi \cdot (a^2 - \frac{3}{2}) \square U_{(1)},$$

which reduces to

$$p_\mu \varphi_{(1)}^\mu = 0, \tag{A.7}$$

because $a^2=3/2$ resulting from the relation (2.27). $U_{(1)}$ thus vanishes in this case. Eliminating $\theta^{\mu\nu}$ and $U_{(1)}$ and using (A.7) one gets

$$(\square - \chi^2) \varphi_{(1)}^\mu = 0. \quad (A \cdot 8)$$

(A·7) and (A·8) show that $\varphi_{(1)}^\mu$ is the μ -th component of a vector potential.

References

- 1) H. J. Bhabha, Phil. Mag. **43** (1952), 33.
- 2) K. K. Gupta, Proc. Roy. Soc. **A 222** (1954), 118.
- 3) H. J. Bhabha, Rev. Mod. Phys. **17** (1945), 200.
- 4) E. Wild, Proc. Roy. Soc. **A 191** (1947), 253.
- 5) H. J. Bhabha, Rev. Mod. Phys. **21** (1949), 451.
- 6) K. J. LeCouteur, Proc. Roy. Soc. **A 202** (1950), 284.
- 7) K. J. LeCouteur, Proc. Roy. Soc. **A 202** (1950), 394.

General Theory of Deuteron Induced Reaction, I

——(d, p) and (d, n) Reactions——

Haruo UI

Institute of Science and Technology, University of Tokyo, Tokyo

(Received April 20, 1956)

Taking into account the loosely bound nature of deuteron, the general theory of compound nucleus formation by deuterons is developed, in which the compound nucleus formation by deuterons is considered as a two step process; first the neutron enters into the target nucleus to form a compound state and then the proton, combined with it, forms another compound state. Stripping process is obtained as the correction terms of the compound nucleus process due to the interaction in the external proton channel. The indirect stripping process is also discussed.

§ 1. Introduction

The theories of deuteron stripping reaction developed by Butler¹⁾ and other authors^{2),3)} have succeeded in explaining the forward peak of the angular distribution of (d, p) reactions in 5—20 Mev energy regions of the incident deuteron and have provided a powerful tool in nuclear spectroscopy to determine the spin and parity of the lowlying states of the nucleus.

These theories, however, seem to fail to explain the angular distribution at larger angles. In the case of $O^{16}(d, p)O^{17}$ reaction by 1–3.5 Mev deuteron,⁴⁾ the angular distributions vary rapidly with a small change of incident deuteron energy. These variations become still more remarkable in kev energy region. Such phenomena seem to occur commonly in the case of light nuclei. On the other hand, the proton energy spectrum, observed at 90° direction through 15 Mev deuteron bombardment⁵⁾ shows the Maxwell distribution characteristic of the statistical property of compound nucleus reaction. These phenomena may hardly be explained by a simple modification of the theory of stripping reaction, e.g. by taking account of the configuration mixing of the states of captured neutron,⁶⁾ the effect of Coulomb interaction on the emitted proton,³⁾ or the polarization of the incident deuteron and the proton-nucleus interaction in the final state.³⁾ The characteristic deviations from the current theory of deuteron stripping reaction mentioned above, therefore, may be ascribed to the interference effects⁷⁾ of the stripping and the compound nucleus processes. Thus it is required to formulate the theory of deuteron induced reaction so as to be capable of treating the stripping and compound nucleus processes on the equal footing.

Since the deuteron is a loosely bound system, having a diameter comparable to or larger than those of light nuclei, it may not be sufficient to match smoothly, on the

boundary surface, the wave function of the motion of the centre of mass of the deuteron, while it is for the case of an elementary particle such as a neutron. In other words, the deuteron enters into a nucleus to form a compound nucleus not only through the so-called deuteron channel but also through the neutron and proton channels. It can be inferred from the fact that the stripping reaction is the main process in (d, p) reaction and that the large cross section of deuteron inelastic scattering can not be explained by the ordinary compound nucleus theory. In fact, the stripping process may be interpreted as follows; the neutron which constitutes the incident deuteron in the deuteron channel enters into the target nucleus through the neutron channel and the proton escapes through the proton channel without entering into the internal region of the residual nucleus. In the formal theory of the nuclear reaction by Teichman and Wigner² the channel wave function, $\psi_s^i, \psi_s^c, \dots$ are defined so as to form an orthonormal set on the surface S , $\int_S \psi_s^{i*} \psi_s^c dS = \delta_{si}$. On the other hand, a solution of the Schrödinger equation in the external region is given by

$$\sum_s \psi_s^i(i_s, Q_s) [c_s F_s(r_s) + c'_s G(r_s)],$$

with arbitrary coefficients of c_s, c'_s , where $F_s(r_s)$ and $G_s(r_s)$ represent the so-called radial wave functions of the pair s , such as a target nucleus and a deuteron. Furthermore, the channel wave function ψ_s^c is assumed to be identified with the wave function of the reaction product ψ_s^r ; $\psi_s^c = \psi_s^r$. As is mentioned above, this identification may no longer hold in the case of the reaction of a loosely bound system, provided the surface S is located near the actual nuclear surface as have been done so far.

To avoid this difficulty, there might be several possible ways. First, the surface S , which divides the configuration space into the internal and external regions, may be located so distant from the actual nuclear surface that not only the compound nucleus process but also the stripping process may take place in the internal region. Then, the stripping process may also be contained in the pole expansion of the R -matrix, $R = \sum_{\lambda} (E - E_{\lambda})^{-1} \cdot \gamma_{\lambda} \times \gamma_{\lambda}$. However, the signs of the $\gamma_{\lambda i}$ of various levels and channels are then not at random but correlated with one another. The statistical theory can not be applied to such a system. The angular distribution of (d, p) reaction derived by such a theoretical formulation could explain the experimental data not only at small angles but also at large ones. This formulation, however, may be of no practical use, because the actual calculations of the deuteron reduced width $\gamma_{\lambda d}$ might be almost impossible in the frame of the mentioned formulation. An alternative way to the above is to construct the theory in such a way as to match smoothly not only the wave function in the deuteron channel but also its overlapping regions in the neutron and proton channels, locating the surface S in the same place as in the case of the current theory. The parameters which appear in this formalism, such as the widths, the positions of the levels, etc., will have the same values as in the current theory.

In this paper we intend to treat exactly the overlapping regions of the wave function in the neutron and proton channels, by neglecting contribution from ordinary deuteron channel. Thus, intuitively speaking, the compound nucleus formation by deuterons is

considered as a two step process in which first the neutron enters into the target nucleus to form a compound state II and then the proton combined with the state II forms the compound state IV and vice versa. For an illustration of our procedure, consider the simplified one-dimensional model used by Butler which is shown in Figure 1. All wave functions are connected smoothly in the regions I, II, III and IV, respectively. As suggested by Thomas,⁷⁾ we can not define the external region in such a way that there are no nuclear interactions in this region, since an incident particle does interact with the tail of the wave function of a target nucleus which extends to infinity, though in an exponentially decaying manner. It is this external region interaction between the proton and the intermediate state II in the region II that causes the stripping process for (d, p) reaction.*

In section 2 a method is described to derive the nuclear reaction theory of Kapur and Peierls⁹⁾ by means of the operator calculus developed by Lippman and Schwinger¹⁰⁾. In Kapur and Peierls', T -matrix is represented directly as a pole expansion of the energies of the levels. On the other hand, R -matrix is given as a pole expansion in the theory of Wigner and Eisenbud.¹¹⁾ The transformation of R -matrix into T -matrix is very complicated, when many levels are contained in R -matrix. Both theories, however, are completely equivalent as is proved by Kawai and Nagasaki.¹²⁾ As it becomes apparent later on, many levels are actually concerned in our formulation. This is the reason why we adopt the theory of Kapur and Peierls. The exact formula of the correction term due to the interactions in the external region is also presented in section 2. In section 3 T -matrix of (d, p) reaction is represented in the form of the sum of two terms; One corresponds to the stripping reaction and the other to the compound nucleus process. By making use of Born approximation the former can be identified with the formula of Butler. The latter is represented, using the method of section 2, in the form of the pole expansion in section 4. In section 5 the term which may be interpreted as a correction of the stripping process due to a certain sort of volume interactions is expressed explicitly. The indirect stripping processes are also discussed.

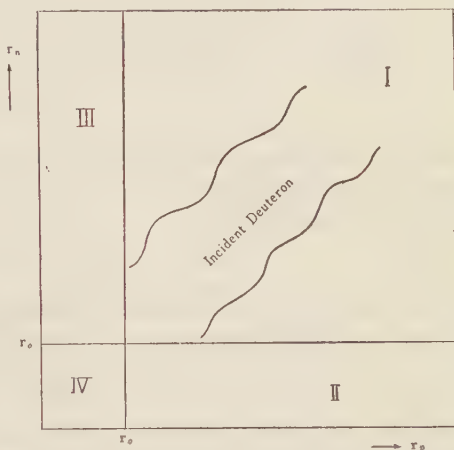


Fig. 1 Simplified one dimensional model of Butler. Region I represents the incident and elastically scattered deuteron, II the neutron in the internal and proton in the external regions, III the proton in the internal and neutron in the external regions and IV compound nucleus by the deuteron capture.

* If we neglect this interaction, the stripping process can not be obtained by means of matching the wave function only between I and II, which was proved by Gerjouy²⁾. See, also section 3.

§ 2. General formulation of nuclear reaction

i) In the first part of this section we shall formulate the theory of nuclear reaction derived by Kapur and Peierls by using the operator calculus of Lippman and Schwinger. The treatment presented here is almost the same as that of Hayakawa.¹³⁾

Consider the following type of reaction,

$$a + A \rightarrow b + B, \quad (1)$$

where a and b are the incident and emitted particles, and A and B are the target and residual nuclei, respectively. The Hamiltonian, H , of the total system is separated into two terms. In the initial state,

$$H = H_a^0 + V_a, \quad (2a)$$

where H_a^0 is the free Hamiltonian, representing the particle a and the nucleus A , and V_a is the interaction between them. In the same manner, for the final state, we have

$$H = H_b^0 + V_b, \quad (2b)$$

in which H_b^0 and V_b are the corresponding quantities in the final state. The wave functions of the free systems in the initial and final states, ϕ_a and ϕ_b , obey the following equations;

$$H_a^0 \phi_a = E \phi_a, \quad H_b^0 \phi_b = E \phi_b. \quad (3)$$

The wave function of the whole system, which is subjected to

$$H\psi = E\psi, \quad (4)$$

is represented, with suitable boundary conditions, in the integral form as follows,

$$\begin{aligned} \psi_a^{(+)} &= \phi_a + (E + i\epsilon - H_a^0)^{-1} V_a \psi_a^{(+)} \\ &= \phi_a + (E + i\epsilon - H)^{-1} V_a \phi_a. \end{aligned} \quad (5)$$

The matrix element of the transition matrix between the initial state ϕ_a and the final state ϕ_b is given by

$$T_{ba} = \langle \phi_b | V_b | \psi_a^{(+)} \rangle = \langle \psi_b^{(-)} | V_a | \phi_a \rangle, \quad (6)$$

where

$$\psi_b^{(-)} = \phi_b + (E - i\epsilon - H)^{-1} V_b \phi_b.$$

In order to derive the Kapur and Peierls' formula, we must expand the wave function $\psi_a^{(+)}$ by the complete set of the decaying states in the internal region. As $\psi_a^{(+)}$ does not satisfy the boundary condition of the decaying state, it can not be expanded as a whole. To avoid the mentioned difficulty, Kapur and Peierls introduced an arbitrary function χ which only satisfies the boundary condition of the incident wave on the surface S . Thus, $\psi_a^{(+)} - \chi$ can be expanded as a whole. As an alternative, we introduce* a potential

* It is found necessary for obtaining (14) to use our procedure of introducing the potential L and making it a hard sphere. This is due to some mathematical reason relating to the convergency of the series (11) on the boundary surface S .

L_a in the initial state, i.e, the interaction is divided as $V_a = (V_a - L_a) + L_a$. To reserve the symmetric property of S -matrix explicitly, it is convenient to divide also the interaction of the final state; $V_b = (V_b - L_b) + L_b$. Then, the transition matrix element becomes as¹⁴⁾

$$\begin{aligned} T_{ba} &= \langle \phi_b | L_b | g_a^{(+)} \rangle + \langle g_b^{(-)} | V_b - L_b | \psi_a^{(+)} \rangle \\ &= \langle \phi_b | L_b | g_a^{(+)} \rangle + \langle g_b^{(-)} | (V_b - L_b) + (V_b - L_b) (1/E + i\epsilon - H) (V_a - L_a) | g_a^{(+)} \rangle, \end{aligned} \quad (7)$$

where

$$g_a^{(+)} = \phi_a + (E + i\epsilon - H_a)^{-1} L_a g_a^{(+)}, \quad (8a)$$

and

$$g_b^{(-)} = \phi_b + (E - i\epsilon - H_b) L_b^{-1} g_b^{(-)}. \quad (8b)$$

The first term of (7), which appears only for $a=b$, represents the scattering by the potential L_a .

Next, we define a set of eigenfunctions, $|X_\lambda\rangle$, of H in the internal region, by

$$H|X_\lambda\rangle = W_\lambda|X_\lambda\rangle, \quad (9a)$$

with the boundary condition of the decaying states* on the surface S . As is well known, such a boundary condition is not Hermitian. Therefore, the eigenvalues are complex and the set $|X_\lambda\rangle$ can not form an orthonormal set by itself. To obtain the orthogonal set of the decaying states, we must introduce another set of the eigenfunctions $|Y_\mu\rangle$ of H with the boundary condition which is complex conjugate to that of $|X_\lambda\rangle$, through

$$H|Y_\mu\rangle = W'_\mu|Y_\mu\rangle. \quad (9b)$$

Since H is a Hermitian operator, we can easily obtain the following relations,

$$\bar{W}'_\lambda = W_\lambda, \quad (\bar{W} \text{ means complex conjugate of } W) \quad (9c)$$

$$\langle Y_\mu | X_\lambda \rangle = \delta_{\lambda\mu} N_\lambda, \quad (10a)$$

and

$$\sum |X_\lambda\rangle (1/N_\lambda) \langle Y_\lambda| = 1, \quad (10b)$$

with a normalization

$$|\langle X_\lambda | X_\lambda \rangle|^2 = 1.$$

These mean that the sets of the eigen functions $|X_\lambda\rangle$ and $\langle Y_\lambda|$ form a biorthogonal set.¹⁵⁾

As the Green function $1/(E + i\epsilon - H)$ obeys the same boundary condition as $|X_\lambda\rangle$, we can expand T-matrix by using (10b),

$$\begin{aligned} &\langle g_b^{(-)} | (V_b - L_b) + (V_b - L_b) \frac{1}{E + i\epsilon - H} (V_a - L_a) | g_a^{(+)} \rangle \\ &= \langle g_b^{(-)} | V_b - L_b | g_a^{(+)} \rangle + \sum_\lambda \langle g_b^{(-)} | V_b - L_b | X_\lambda \rangle \frac{1}{(E - W_\lambda) N_\lambda} \langle Y_\lambda | V_a - L_a | g_a^{(+)} \rangle, \end{aligned} \quad (11)$$

* See, Eq. (65) of reference 9.

where a small imaginary quantity $+i\epsilon$ is omitted, for brevity, in the denominator of the last term. Using the following identities derived from the Schrödinger equations satisfied by $g_a^{(+)}$ and $g_b^{(-)}$, and $|X_\lambda\rangle$ and $\langle Y_\lambda|$,

$$\langle Y_\lambda | V_a - L_a | g_a^{(+)} \rangle = \gamma_{\lambda a} - (E - W_\lambda) \langle Y_\lambda | g_a^{(+)} \rangle, \quad (12a)$$

$$\langle g_b^{(-)} | V_b - L_b | X_\lambda \rangle = \gamma_{b\lambda} - (E - W_\lambda) \langle g_b^{(-)} | X_\lambda \rangle, \quad (13b)$$

$$\begin{aligned} \gamma_{\lambda a} &\equiv \langle Y_\lambda | H | g_a^{(+)} \rangle - \langle Y_\lambda | \tilde{H} | g_a^{(+)} \rangle, \\ \gamma_{b\lambda} &\equiv \langle g_b^{(-)} | \tilde{H} | X_\lambda \rangle - \langle g_b^{(-)} | H | X_\lambda \rangle, \end{aligned} \quad (\tilde{H} = \text{transpose of } H)$$

the right hand side of (11) can be reduced to

$$\sum_\lambda \gamma_{b\lambda} \frac{1}{(E - W_\lambda) N_\lambda} \gamma_{\lambda a} - \sum_\lambda \langle g_b^{(-)} | X_\lambda \rangle \frac{1}{N_\lambda} \gamma_{\lambda a} - \sum_\lambda \gamma_{b\lambda} \frac{1}{N_\lambda} \langle Y_\lambda | g_a^{(+)} \rangle. \quad (13)$$

The second and third terms of (13) become, by making use of the relation (10b),

$$\langle g_b^{(-)} | H | g_a^{(+)} \rangle - \langle g_b^{(-)} | \tilde{H} | g_a^{(+)} \rangle \quad \text{and} \quad \langle g_b^{(-)} | \tilde{H} | g_a^{(+)} \rangle - \langle g_b^{(-)} | H | g_a^{(+)} \rangle.$$

Therefore, the matrix element of T-matrix between the states a and b is given by

$$T_{ba} = \langle \phi_b | L_b | g_a^{(+)} \rangle \delta_{ab} + \sum_\lambda \gamma_{b\lambda} \{ (E - W_\lambda) N_\lambda \}^{-1} \gamma_{\lambda a}. \quad (14)$$

Taking the potential L_a and L_b as a hard sphere on the surface S , the formula (14) can be proved to agree with that of Kapur and Peierls exactly. Our procedure for L_a and L_b above just corresponds to the situation where in the latter formalism a function χ was introduced with the condition that it is zero in everywhere in the internal region and has a certain value of derivative only on the surface S .

Now, there is a convenient relation for the calculation of the (amplitude of) width,

$$\begin{aligned} \gamma_{\lambda a} &= \langle Y_\lambda | H | g_a^{(+)} \rangle - \langle Y_\lambda | \tilde{H} | g_a^{(+)} \rangle \\ &= \langle Y_\lambda | H | \phi_a \rangle - \langle Y_\lambda | \tilde{H} | \phi_a \rangle. \end{aligned} \quad (15)$$

By the use of Green's theorem the first expression of (15) can be transformed into a surface integral as

$$\langle Y_\lambda | H | g_a^{(+)} \rangle - \langle Y_\lambda | \tilde{H} | g_a^{(+)} \rangle = \frac{\hbar^2}{2M} \int_S \left\{ Y_\lambda^* \frac{\partial}{\partial r} g_a^{(+)} - \frac{\partial}{\partial r} Y_\lambda^* \cdot g_a^{(+)} \right\} dS,$$

which is readily transformed to the final expression (15), remembering that $g_a^{(+)}$ is, on the surface S , ϕ_a plus outgoing wave with a proportional factor and that the boundary condition imposed on Y_λ^* is identical with that of the outgoing wave.

In a similar way as above, the R -matrix of Wigner and Eisenbud can be derived by taking a self-adjoint boundary condition and the corresponding orthonormal set of eigenfunctions in which the term corresponding to the first term of (14) does not appear, as is apparent by the definition of R -matrix.

ii) Surface direct interaction—Interaction in the external region

In the formal theory of nuclear reaction it has been assumed that all nuclear interactions are confined to the internal region of the configuration space. In order to apply the statistical treatment, the surface S should be located, as far as possible, near the actual nuclear surface. However, an incident particle does interact with the tail of the wave function of a target nucleus which extends to infinity, though in an exponentially decaying manner. Therefore, we should take into account the contribution of the interactions in the external region, especially when the S is located near the nuclear surface.

For this purpose, we divide the interaction V into two parts, $V = V^I + V^E$, V^I is the interaction in the internal region and V^E that in the external one. For the initial state a , $V_a = V_a^I + V_a^E$, and for the final state b , $V_b = V_b^I + V_b^E$. In the same manner as (7), we have

$$T_{ba} = \langle \phi_b | V_b^I | G_a^{(+)} \rangle + \left\langle G_b^{(-)} | V_b^E \left[1 + \frac{1}{E + i\epsilon - H} V_a^E \right] G_a^{(+)} \right\rangle \quad (16)$$

with

$$G_a^{(+)} = \left[1 + \frac{1}{E + i\epsilon - H_a^0 - V_a^I} V_a^I \right] \phi_a,$$

$$G_b^{(-)} = \left[1 + \frac{1}{E - i\epsilon - H_b^0 - V_b^I} V_b^I \right] \phi_b.$$

The first term represents the contributions from the internal region only and is identical with the eq. (14), since in (14) the interactions in the external region are completely neglected. The second term is the correction term due to the interactions in the external region. The interaction V^E may be treated as a small perturbation, because the tail of the wave function in the external region may be considered to be very small. Then, (16) becomes approximately

$$T_{ba} \simeq \langle \phi_b | V_b^I | G_a^{(+)} \rangle + \langle G_b^{(-)} | V_b^E | G_a^{(+)} \rangle, \quad (17)$$

the second term of which is nothing but the modified (distorted wave) Born approximation.

§ 3. T-matrix of (d, p) reaction and stripping process

First we shall prepare the formula of T-matrix in the most suitable form for treating our problem. The Hamiltonian H of the system is given by

$$H = H_N + T_p + T_n + V_{np} + V_p + V_n, \quad (18)$$

where H_N is the Hamiltonian of the target nucleus, T_p and T_n are the kinetic energy operators of the proton and neutron, and V_{np} , V_p and V_n the interaction between the proton and neutron, that between the proton and target nucleus, and that between the neutron and target nucleus, respectively. In the initial state in which we have the incident deuteron and the target nucleus the Hamiltonian of free state is $H_i = H_N + T_n + T_p + V_{np}$ and the interaction Hamiltonian is taken to be $V_i = V_n + V_p$, whereas in the final state

of the residual nucleus and the emitted proton, the corresponding quantities are given by $H_f = (H_N + T_n + V_n) + T_p$ and $V_f = V_p + V_{np}$, respectively. The wave function of the total system can be written as

$$\Psi_i^{(+)} = \left[1 + \frac{1}{E + i\epsilon - H} (V_n + V_p) \right] \Phi_i, \quad (19)$$

with

$$(H_N + T_n + T_p + V_{np}) \Phi_i = E \Phi_i. \quad (20a)$$

Φ_i is the free state wave function in the initial state which is expressed by the product of the wave functions of the incident deuteron and the target nucleus.

Then, the T -matrix element is given by

$$T_{pi} = \langle \Phi_f | V_p + V_{np} | \Psi_i^{(+)} \rangle, \quad (21)$$

with

$$\{ (H_N + T_n + V_n) + T_p \} \Phi_f = E \Phi_f, \quad (20b)$$

where Φ_f is a free state wave function of the final state, i.e., the product of the wave functions of the proton plane wave and of the residual nucleus. The wave function $\Psi_i^{(+)}$ which is given by (19) can be written as follows;

$$\begin{aligned} \Psi_i^{(+)} = & \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_{np} - V_n - V_p} (V_p + V_{np}) \right] \\ & \times \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_n} (V_n - V_{np}) \right] \Phi_i, \end{aligned} \quad (22)$$

which can be easily obtained by algebraic manipulation. This relation is more convenient for our problem than the corresponding expression used by Francis and Watson,¹⁶⁾ i.e.,

$$\begin{aligned} \Psi_i^{(+)} = & \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_{np} - V_n - V_p} V_p \right] \\ & \times \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_{np} - V_n} V_n \right] \Phi_i, \end{aligned}$$

for the reason that the denominator in the second brackets does not contain the interaction operator V_{np} , and the interaction appeared in the first brackets is $V_p + V_{np}$ rather than V_p . The former may be required mainly from the mathematical point of view, which will become apparent in the next section, whereas the latter is due to the physical reason that the interaction $V_p + V_{np}$ on the proton can not be considered separately after the neutron is captured in the nucleus. Moreover, $\varphi = [1 + (E + i\epsilon - H_N - T_n - T_p - V_n)^{-1} \times (V_n - V_{np})] \Phi_i$ is one of the wave functions in the region II of Fig. 1., since it satisfies the following Schrödinger equation,

$$[(H_N + T_n + V_n) + T_p] \varphi = E \varphi.$$

It is to be noted that the wave matrix $[1 + (E + i\epsilon - H_N - T_n - T_p - V_n)^{-1} \cdot (V_n - V_{np})]$

does not create an outgoing wave of proton, because the amplitude of the proton outgoing wave in φ is proportional to $\{\langle \phi_f | V_n | \phi_i \rangle - \langle \phi_f | V_{np} | \phi_i \rangle\}$, which is apparently zero.

As is discussed in Introduction and section 2, we must take into account the interaction between the proton and the tails of the wave function of the state which is formed by the neutron capture into the target nucleus. This interaction is denoted by $(V_p + V_{np})^{II}$; $V_p + V_{np} = (V_p + V_{np})^{II} + (V_p + V_{np})^{IV}$, where $(V_p + V_{np})^{IV}$ is the interaction acting on the proton in the region IV. Then, (21) can be written, in the same manner as (16), as follows;

$$T_{pd} = (\phi_f | V_p + V_{np})^{IV} | \mathcal{Q}_{IV}^{(+)} \phi_i \rangle + \left(G_f^{(-)} | (V_p + V_{np})^{II} \left[1 + \frac{1}{E + i\epsilon - H} (V_p + V_{np})^{II} \right] \mathcal{Q}_{IV}^{(+)} \phi_i \right) \quad (23)$$

$$\equiv T_{pd}^{(c')} + T_{pd}^{(s)},$$

where

$$G_f^{(-)} = \left[1 + \frac{1}{E - i\epsilon - H_{II} - T_p - (V_p + V_{np})^{IV}} (V_p + V_{np})^{IV} \right] \phi_f,$$

and

$$\mathcal{Q}_{IV}^{(+)} = \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - (V_n + V_{np})^{IV}} (V_p + V_{np})^{IV} \right] \\ \times \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p} (V_n + V_{np}) \right],$$

$$H_{II} \equiv H_N + T_n + V_n.$$

The first term represents the contributions only from the region IV and may be interpreted as a compound nucleus process, which will be treated in the section 4. The second is the correction term due to the interaction with the proton in the region II and may be interpreted as the stripping process arising from the so-called surface interaction.

As $(V_p + V_{np})^{II}$ is the interaction between the tails of the wave function of the intermediate states formed by the neutron capture and the proton in the external region, it may be expected to be small compared with the volume interaction $(V_p + V_{np})^{IV}$. Therefore, we may treat it as a perturbation. Then the second term of (23) can be written as

$$T_{pd}^{(s)} \simeq (G_f^{(-)} | (V_p + V_{np})^{II} | \mathcal{Q}_{IV}^{(+)} \phi_i). \quad (24)$$

The neutron in the incident deuteron is likely to be captured in an outer orbit; for example, in the Fermi gas model the neutron is captured into a level outside of the Fermi surface with the residual nucleus in its ground state. Therefore, V_p^{II} may be small compared with V_{np}^{II} , since the wave function of the captured neutron has a longer tail than those of the other nucleons in the nucleus. If V_p^{II} is neglected in (24), we have

$$T_{pd}^{(s)} \simeq (G_f^{(-)} | V_{np}^{II} | \mathcal{Q}_{IV}^{(+)} \phi_i). \quad (25)$$

Since the interactions contained in $G_f^{(-)}$ and $\mathcal{Q}_{IV}^{(+)} \phi_i$ are confined to the region IV, $G_f^{(-)}$

and $\mathcal{Q}_{fV}^{(+)}\phi_i$ in (24) and (25) may be written in the asymptotic forms as follows;

$$G_f^{(-)} = (\text{time reversed state of}) \quad \phi_f + \sum_{\alpha} T_{\alpha f} E_{\alpha}$$

$$\mathcal{Q}_{fV}^{(+)}\phi_i = \phi_i + \sum_{\beta} T_{\beta i} E_{\beta}$$

where E_{α} and E_{β} are out-going waves in the channel α and β , respectively. If we take approximately only the waves in the channel i and f , (25) becomes the formula of the modified (distorted) Born approximation of Tobočan, which was proved by Thomas.⁷⁾

For the analysis of the experimental data it seems reasonable to take the waves distorted by the complex potential which is determined by the experiments of the elastic scattering. For $G_f^{(-)}$, we may safely use the complex potential of Feshbach, Porter and Weisskopf¹⁷⁾ which is determined by the scattering experiments of the proton by the residual nucleus χ_f , when the residual nucleus χ_i is in the ground state. If it is not in the ground state, it may be impossible to determine the optical potential empirically. At least, it may be supposed that the complex potential has a larger imaginary part than that of the ground state.¹⁸⁾ On the other hand, the experimental data of deuteron elastic scattering are rather scanty to determine the distorted wave in the entrance channel of deuteron. The accumulation of the data is wanted.

If we use Born approximation in (25), the formula of Butler can be obtained.*

§ 4. Compound nucleus formation by deuterons

As is discussed in Introduction, we consider the compound nucleus formation by deuterons as a two step process in which first the neutron enters into the target nucleus to form the compound state II and then the proton, combined with it, forms the compound state IV and vice versa. Accordingly, the configuration space is divided into four parts; the external region I in the ordinary sense, region II where the neutron is in the internal region and the proton in the external region, region III where the roles of the neutron and proton are exchanged, and the internal region IV in the ordinary sense. The surfaces between each of them are called S^{12} , S^{13} , S^{24} and S^{34} , respectively. In this section we perform the pole expansion of the first term of (22), which represents the process through the region IV. Throughout this section the suffix IV of $(V_p + V_{np})^{IV}$ is omitted for brevity.

For expanding T -matrix in terms of the biorthogonal set of decaying states, we must introduce the potentials L_n and L_p in the same manner as in section 2;

$$V_n = (V_n - L_n) + L_n,$$

and

$$V_p + V_{np} = (V_p + V_{np} - L_p) + L_p.$$

It must be noted that the potential L_p is in general not diagonal for the states of the

* See the Appendix of reference 7.

target nucleus in contradiction with the case of L_n , but is taken to be diagonal for the states of the residual nucleus. In the analogous way as in (7) of section 2, we obtain

$$\mathcal{Q}_{IV}^{(+)} = \left[1 + \frac{1}{E + i\epsilon - H} (V_n + V_{np} - L_p) \right] \prod \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_n - L_p} L_p \right] \\ \times \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_n} (V_n - V_{np} - L_n) \right] \prod \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_{np} - L_n} L_n \right],$$

on account of which the T -matrix of the process given by the first term of (23) can be written as

$$\langle \phi_f | V_n + V_{np} | \mathcal{Q}_{IV}^{(+)} \phi_i \rangle \\ = \langle \phi_f | L_p \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - L_p} L_p \right] \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_n} (V_n - V_{np}) \right] | \phi_i \rangle \\ + \langle g_f^{(-)} | V_p + V_{np} - L_n | \mathcal{Q}_{IV}^{(+)} \phi_i \rangle \equiv T_{pd}^{(s')} + T_{pd}^{(c)} \quad (26)$$

with

$$g_f^{(-)} = \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - L_p} L_p \right] \phi_f. \quad (27)$$

The first term may be interpreted as follows; the neutron is captured by the target nucleus, which is caused by the wave matrix, $1 + (E + i\epsilon - H_{II} - T_p)^{-1} \cdot (V_n - V_{np})$, and the proton is scattered by the potential L_p , without changing the state of the nucleus which is formed by the neutron capture. This process will be discussed in detail in the next section. The second term represents the process through the so-called compound nucleus formation by deuteron as a whole as will be apparent later. The following calculations are divided into two parts for convenience; in the part (i) the pole expansion of the T -matrix in the region II is performed and in the part (ii) the final expression of the pole expansion is presented.

(i) Pole expansion in the region II.

We now expand the T -matrix in the region II where the neutron is in the internal region and the proton in the external region in the configuration space.

Before the expansion is performed, we must separate the coordinates of the neutron and proton. For this purpose we define the complete set of the free states of the proton wave,

$$T_p | \mathbf{k}_p \rangle = \varepsilon_p | \mathbf{k}_p \rangle, \quad \varepsilon_p = \hbar^2 k_p^2 / 2M_p.$$

Then

$$\phi_i = \sum_{\mathbf{k}_p} | \mathbf{k}_p \rangle \langle \mathbf{k}_p | \phi_i \rangle \equiv \sum_{\mathbf{k}_p} | \mathbf{k}_p \rangle \cdot g_n(\mathbf{k}_p, \mathbf{r}_n, \xi)$$

Here ξ means all coordinates of the target nucleus. From the Schrödinger equation satisfied by ϕ_i , i.e.,

$$(E - H_N - T_n - T_p) \phi_i = V_{np} \phi_i,$$

* In this section we use the notation H , for simplicity, instead of $H - (V_p + V_{np})^{II}$.

it follows that

$$(E'' - H_N - T_n) g'_n(\mathbf{k}_p, \mathbf{r}_n, \xi) = b'_n(\mathbf{k}_p, \mathbf{r}_n, \xi),$$

where

$$b'_n(\mathbf{k}_p, \mathbf{r}_n, \xi) = \langle \mathbf{k}_p | V_{np} \phi_i \rangle,$$

and

$$E'' = E - \varepsilon_p.$$

Therefore,

$$\begin{aligned} \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p} (V_n - V_{np}) \right] \phi_i &= \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p} (V_n - V_{np} - L_n) \right] G^{(+)} \\ &= \sum_{\mathbf{k}_p} |\mathbf{k}_p\rangle \cdot \left[g_n^{(+)} + \frac{1}{E'' + i\epsilon - H_{II}} \{ (V_n - V_{np}) - b_n^{(+)} \} \right], \end{aligned} \quad (28)$$

where

$$G^{(+)} = \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_{np} - L_n} L_n \right] \phi_i, \quad (29)$$

and

$$g_n^{(+)}(\mathbf{k}_p, \mathbf{r}_n, \xi) = \langle \mathbf{k}_p | G^{(+)} \rangle, \quad (30a)$$

$$b_n^{(+)}(\mathbf{k}_p, \mathbf{r}_n, \xi) = \langle \mathbf{k}_p | V_{np} G^{(+)} \rangle. \quad (30b)$$

We next define the biorthogonal set, $|II_\lambda\rangle$ and $\langle II'_\lambda|$, of the decaying states of the collision complex formed by the neutron and target nucleus,

$$H_{II} |II_\lambda\rangle = W_\lambda^{II} |II_\lambda\rangle, \quad (31a)$$

with the boundary condition of the decaying state on the surface S^{in} , and

$$H_{II} |II'_\lambda\rangle = \bar{W}_\lambda^{II} |II'_\lambda\rangle, \quad (31b)$$

with the complex conjugate boundary condition to that of $|II_\lambda\rangle$. Then, it follows that

$$\sum_\lambda |II_\lambda\rangle \langle II'_\lambda| = 1 \quad (32)$$

with the normalization

$$|\langle II_\lambda | II_\lambda \rangle|^2 = 1, \quad \langle II'_\lambda | II'_\mu \rangle = N_\lambda^{II} \delta_{\lambda\mu}.$$

Hence, we have, by the aid of (32)

$$\begin{aligned} &\left[g_n^{(+)} + \frac{1}{E + i\epsilon - H_{II}} \{ (V_n - L_n) g_n^{(+)} - b_n^{(+)} \} \right] \\ &= g_n^{(+)} + \sum_\lambda |II_\lambda\rangle \frac{1}{(E'' - W_\lambda^{II}) N_\lambda^{II}} \{ \langle II'_\lambda | V_n - L_n | g_n^{(+)} \rangle - \langle II'_\lambda | g_n^{(+)} \rangle \}. \end{aligned} \quad (34)$$

Now, we have the relation

$$\langle II'_\lambda | V_n - L_n | g_n^{(+)} \rangle - \langle II'_\lambda | b_n^{(+)} \rangle = \omega_\lambda^{II} - (E'' - W_\lambda^{II}) \langle II'_\lambda | g_n^{(+)} \rangle, \quad (35)$$

with

$$\omega_{\lambda n}^{II} = \langle II_{\lambda}^c | H_{II} | g_n^{(+)} \rangle = \langle II_{\lambda}^c | \tilde{H}_{II} | g_n^{(+)} \rangle,$$

which can be obtained by the following Schrödinger equations

$$(H_{II} - E^{II}) g_n^{(+)} = (V_n - L_n) g_n^{(+)} - h_n^{(+)},$$

and

$$\langle II_{\lambda}^c | \tilde{H}_{II} = \langle II_{\lambda}^c | W_{\lambda}^{II}.$$

The substitution of Eq (35) in Eq (34) leads to

$$\sum_{\lambda} |II_{\lambda}\rangle \frac{1}{(E^{II} - W_{\lambda}^{II}) N_{\lambda}^{II}} \omega_{\lambda n}^{II}. \quad (36)$$

Therefore, we have

$$\begin{aligned} & \left[1 + \frac{1}{E + i\epsilon - H_N - T_n - T_p - V_n} (V_n - V_{np}) \right] \Phi_i \\ &= \sum_{\mathbf{k}_p} \sum_{\lambda} |\mathbf{k}_p\rangle \cdot |II_{\lambda}\rangle \frac{1}{(E^{II} - W_{\lambda}^{II}) N_{\lambda}^{II}} \omega_{\lambda n}^{II}, \end{aligned} \quad (37)$$

which is the result of the expansion in the region II. The $\omega_{\lambda n}^{II}$ is a factor proportional to the neutron width $\gamma_{\lambda n}^{II}$, the explicit form of which will be given in the next section.

(ii) Pole expansion in the region IV.

In order to perform the pole expansion of the T-matrix in the region IV, we must introduce the biorthogonal set, $|IV_{\mu}\rangle$ and $\langle IV_{\mu}^c|$, of the decaying states of the collision complex which is formed by the target nucleus and the deuteron in the same manner as in the region II.

$$H|IV_{\mu}\rangle = W_{\mu}^{IV} |IV_{\mu}\rangle, \quad (38a)$$

with the boundary condition of the decaying state on the surface S^{04} , and

$$H|IV_{\mu}^c\rangle = \bar{W}_{\mu}^{IV} |IV_{\mu}^c\rangle, \quad (38b)$$

with the complex conjugate boundary condition to the former. Then, we have as before

$$\langle IV_{\mu}^c | IV_{\mu} \rangle = N_{\mu}^{IV}, \quad |\langle IV_{\mu} | IV_{\mu} \rangle|^2 = 1,$$

and

$$\sum_{\mu} |IV_{\mu}\rangle \frac{1}{N_{\mu}^{IV}} \langle IV_{\mu}^c| = 1. \quad (39)$$

Our task in this subsection is to express the following formula, the second term of (26) in the form of pole expansion;

$$\begin{aligned} & \langle g_f^{(-)} | V_p + V_{np} - L_p | \left[1 + \frac{1}{E + i\epsilon - H} (V_p + V_{np} - L_p) \right] \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_{\mu} - L_p} L_p \right] \\ & \quad \times \left[\sum_{\mathbf{k}_p, \lambda} |\mathbf{k}_p\rangle \cdot |II_{\lambda}\rangle \frac{1}{(E^{II} - W_{\lambda}^{II}) N_{\lambda}^{II}} \omega_{\lambda n}^{II} \right] \end{aligned}$$

$$= \langle g_f^{(-)} | V_p + V_{np} - L_p \left[1 + \frac{1}{E + i\epsilon - H} (V_p + V_{np} - L_p) \right] \left[\sum_{\mathbf{k}_p, \lambda} |g_p^{(+)}\rangle \cdot |II_\lambda\rangle \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} \right], \quad (40)$$

where

$$|g_p^{(+)}(\mathbf{k}_p, \mathbf{r}_p)\rangle = \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - L_p} L_p \right] |\mathbf{k}_p\rangle. \quad (41)$$

Allowing for $|g_p^{(+)}\rangle |II_\lambda\rangle = |g_p^{(+)} II_\lambda\rangle$, the expression,

$$\left[1 + \frac{1}{E + i\epsilon - H} (V_p + V_{np} - L_p) \right] |g_p^{(+)} II_\lambda\rangle \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} \quad (42)$$

can be reduced, using (39), to

$$|g_p^{(+)} II_\lambda\rangle \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} + \sum_\mu |IV_\mu\rangle \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \langle IV_\mu^c | V_p + V_{np} - L_p | g_p^{(+)} II_\lambda \rangle \times \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II}. \quad (43)$$

By making use of the relation,

$$\langle IV_\mu^c | V_p + V_{np} - L_p | g_p^{(+)} II_\lambda \rangle = \gamma_{\mu, k_p II_\lambda} - (\varepsilon_p + W_\lambda^{II} - W_\mu^{IV}) \langle IV_\mu^c | g_p^{(+)} II_\lambda \rangle, \quad (44)$$

with

$$\gamma_{\mu, k_p II_\lambda} \equiv \langle IV_\mu^c | H | g_p^{(+)} II_\lambda \rangle - \langle IV_\mu^c | \tilde{H} | g_p^{(+)} II_\lambda \rangle,$$

which is derived by the equations

$$(H - W_\lambda^{II} - \varepsilon_p) |g_p^{(+)} II_\lambda\rangle = (V_p + V_{np} - L_p) |g_p^{(+)} II_\lambda\rangle,$$

and

$$\langle IV_\mu^c | \tilde{H} = \langle IV_\mu^c | W_\mu^{IV},$$

Eq. (42) can be written as

$$\sum_\mu |IV_\mu\rangle \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \left\{ \gamma_{\mu, k_p II_\lambda} \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} + \langle IV_\mu^c | g_p^{(+)} II_\lambda \rangle \frac{1}{N_\lambda^{II}} \omega_{\lambda n}^{II} \right\}, \quad (45)$$

in which (39) and the identity of

$$-\frac{1}{E - W_\mu^{IV}} (\varepsilon_p + W_\lambda^{II} - W_\mu^{IV}) \frac{1}{E^{II} - W_\lambda^{II}} = \frac{1}{E - W_\mu^{IV}} - \frac{1}{E^{II} - W_\lambda^{II}} \quad (46)$$

are used.

Our next task is to express, in the analogous style as (14) in section 2, the following expression

$$\sum_{\mu, \lambda, k_p} \langle g_f^{(-)} | V_p + V_{np} - L_p | IV_\mu \rangle \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \left\{ \gamma_{\mu, k_p II_\lambda} \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} + \langle IV_\mu^c | g_p^{(+)} II_\lambda \rangle \frac{1}{N_\lambda^{II}} \omega_{\lambda n}^{II} \right\}. \quad (47)$$

In the same manner as in (12) it follows that

$$\langle g_f^{(-)} | V_p + V_{np} - L_p | IV_\mu \rangle = \gamma_{p\mu} - (E - W_\mu^{IV}) \langle g_f^{(-)} | IV_\mu \rangle, \quad (48)$$

$$\gamma_{p\mu} = \langle g_f^{(-)} | \tilde{H} | IV_\mu \rangle - \langle g_f^{(-)} | H | IV_\mu \rangle,$$

on account of which (47) can be written as

$$\begin{aligned} & \sum_{\mu, \lambda, k_p} \gamma_{p\mu} \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \left\{ \gamma_{\mu, k_p II \lambda} \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} + \langle IV_\mu^c | g_p^{(+)} II_\lambda \rangle \frac{1}{N_\lambda^{II}} \omega_{\lambda n}^{II} \right\} \\ & - \sum_{\mu, \lambda, k_p} \langle g_f^{(-)} | IV_\mu \rangle \frac{1}{N_\mu^{IV}} \gamma_{\mu, k_p, II \lambda} \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} - \sum_{\lambda, k_p} \langle g_f^{(-)} | g_p^{(+)} II_\lambda \rangle \frac{1}{N_\lambda^{II}} \omega_{\lambda n}^{II}. \end{aligned} \quad (49)$$

The last two expressions can be proved to vanish because of the identities,

$$\begin{aligned} \sum_{\mu} \langle g_f^{(-)} | IV_\mu \rangle \frac{1}{N_\mu^{IV}} \gamma_{\mu, k_p, II \lambda} &= \langle g_f^{(-)} | H | g_p^{(+)} II_\lambda \rangle - \langle g_f^{(-)} | \tilde{H} | g_p^{(+)} II_\lambda \rangle \\ &= (E^{II} - W_\lambda^{II}) \langle g_f^{(-)} | g_p^{(+)} II_\lambda \rangle, \end{aligned}$$

and

$$\sum_{\lambda, k_p} \langle g_f^{(-)} | g_p^{(+)} II_\lambda \rangle \frac{1}{N_\lambda^{II}} \omega_{\lambda n}^{II} = \langle g_f^{(-)} | H_{II} | g_p^{(+)} g_n^{(+)} \rangle - \langle g_f^{(-)} | \tilde{H}_{II} | g_p^{(+)} g_n^{(+)} \rangle.$$

In order to make all parameters in our formalism to have the same values as in the current theory, it is necessary to take L_n and L_p as hard sphere potentials. Then the second term in the curled brackets of (49) makes also no contributions.

Then, we have finally, for the T-matrix of compound nucleus reaction,

$$T_{pnl}^{(c)} = \sum_{\mu} \gamma_{p\mu} \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \left\{ \sum_{\lambda, k_p} \gamma_{\mu, k_p II \lambda} \frac{1}{(E^{II} - W_\lambda^{II}) N_\lambda^{II}} \omega_{\lambda n}^{II} \right\}. \quad (50)$$

There is, of course, an alternative process in which first the proton enters into the target nucleus to form the compound state $|III_{\lambda'}\rangle$, and then the neutron forms the compound state $|IV_\mu\rangle$ combined with the state $|III_{\lambda'}\rangle$. T-matrix element of this process is evidently given by

$$\tilde{T}_{pnl}^{(c)} = \sum_{\mu} \gamma_{p\mu} \frac{1}{(E - W_\mu^{IV}) N_\mu^{IV}} \left\{ \sum_{\lambda', k_n} \gamma_{\mu, k_n III \lambda'} \frac{1}{(E^{III} - W_{\lambda'}^{III}) N_{\lambda'}^{III}} \omega_{\lambda' p}^{III} \right\}, \quad (50')$$

where the notations $\gamma_{\mu, k_n III \lambda'}$, $\omega_{\lambda' p}^{III}$, E^{III} and $W_{\lambda'}^{III}$, respectively, are just the corresponding quantities to the $\gamma_{\mu, k_p II \lambda}$, $\omega_{\lambda n}^{II}$, E^{II} and W_λ^{II} in the former process mentioned before, since the roles of the neutron and proton are exchanged. The term corresponding to the first term of (26) vanishes by virtue of the potential L_n' corresponding to L_p , because we are treating only the process in which a proton appears as a reaction product in the final state.

Since the quantity $\gamma_{p\mu}$ is transformed, in the usual way, into a surface integral on the boundary surface, it represents evidently (the amplitude of) the proton width which has appeared in the theory of Kapur and Peierls. On the other hand, the quantity $\gamma_{\mu, k_p II \lambda}$ is the proton width for the process in which the target nucleus is not in the ground state but in an (real or virtual) excited state.

The (amplitude of) deuteron width is reasonably defined as

$$\gamma_{p,d} \equiv \sum_{\lambda, k_p} \tilde{\gamma}_{p, k_p} II_{\lambda} \frac{1}{(E^{II} - W_{\lambda}^{II}) N_{\lambda}^{II}} \omega_{\lambda n}^{II} + \sum_{\lambda', k_n} \tilde{\gamma}_{p, k_n} III_{\lambda'} \frac{1}{(E^{III} - W_{\lambda'}^{III}) N_{\lambda'}^{III}} \omega_{\lambda' p}^{III} \quad (51)$$

Each term in the above summation may be reasonably interpreted as follows; the neutron within the incident deuteron forms a collision complex λ by entering into the target nucleus and then the proton, combined with λ , forms a compound state II , and the alternative process, respectively. In our result obtained above, it may be worth noticing that many compound states λ , intermediately formed by the neutron capture, make actual contributions to the deuteron width which, in some cases, for instance, in the case of indirect stripping process, play an important role, as will be discussed later.

§ 5. The volume direct interaction $T_{pd}^{(s'l)}$

In this section $T_{pd}^{(s'l)}$, the first term of (26), is expressed explicitly. With the aid of the result obtained by (4)-(i), $T_{pd}^{(s'l)}$ can be written as follows:

$$T_{pd}^{(s'l)} = \sum_{k_p, \lambda} \langle \phi_f | L_p \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - L_p} \right] | k_p, II_{\lambda} \rangle \frac{1}{(E^{II} - W_{\lambda}^{II}) N_{\lambda}^{II}} \omega_{\lambda n}^{II} \quad (52)$$

which may be interpreted as follows, the neutron is captured into the target nucleus to form a state $|II_{\lambda}\rangle$, and the proton which is in a free state $|k_p\rangle$ are scattered by L_p without changing the state $|II_{\lambda}\rangle$. Therefore, $T_{pd}^{(s'l)}$ may be considered as a direct process due to volume interaction, whereas the stripping process presented in section 3 is the direct process arising from the surface interaction.

(i) First we shall calculate the $\omega_{\lambda n}^{II}$;

$$\begin{aligned} \omega_{\lambda n}^{II} &= \langle II_{\lambda}^c | H_{II} | g_n^{(+)} \rangle - \langle II_{\lambda}^c | \tilde{H}_{II} | g_n^{(+)} \rangle \\ &= (-\hbar^2/2M_n) \int_{S^{12}} \left\{ \langle II_{\lambda}^c | \frac{\partial}{\partial \mathbf{r}_n} g_n^{(+)} - \frac{\partial}{\partial \mathbf{r}_n} \langle II_{\lambda}^c | \cdot g_n^{(+)} \right\}_{\mathbf{r}_n = \mathbf{r}_0} dS^{12}, \end{aligned} \quad (53)$$

which may be approximately replaced by*

$$\omega_{\lambda n}^{II} = (-\hbar^2/2M_n) \int_{S^{12}} \left\{ \langle II_{\lambda}^c | \frac{\partial}{\partial \mathbf{r}_n} g_n - \frac{\partial}{\partial \mathbf{r}_n} \langle II_{\lambda}^c | g_n \right\}_{\mathbf{r}_n = \mathbf{r}_0} dS^{12}. \quad (54)$$

As was defined in the preceding section, ϕ_i in $g_n = \langle \mathbf{k}_n | \phi_i \rangle$ is the product of the wave functions of the target nucleus $\chi_i(\xi)$ and of the plane wave of the incident deuteron $e^{iK(r_n + r_p)/2} \phi_d(|\mathbf{r}_n - \mathbf{r}_p|)$.

Therefore, we have

$$g_n = \eta(\mathbf{r}_n) \cdot \chi_i(\xi), \quad (55)$$

* In this case Eq. (15) can not hold exactly, because the outgoing wave in $g_n^{(+)}$ does not satisfy the same boundary condition as $\langle II_{\lambda}^c |$.

and

$$\eta(\mathbf{r}_n) = \frac{1}{(2\pi)^{3/2}} \int d\mathbf{r}_p e^{-i\mathbf{k}_p \cdot \mathbf{r}_p} \cdot e^{i\mathbf{K} \cdot (\mathbf{r}_n + \mathbf{r}_p)/2} \cdot \psi_d(|\mathbf{r}_n - \mathbf{r}_p|) \quad (56)$$

where \mathbf{K} is the wave number vector of the centre of mass of the incident deuteron and $\psi_d(|\mathbf{r}_n - \mathbf{r}_p|)$ the wave function of its internal motion.

Taking $\psi_d(r)$ to be of Hulthén type, i.e.,

$$\psi_d(r) = N_d e^{-a/r} (1 - e^{-br}) / r, \quad (57)$$

the integration in (56) is carried through, resulting in

$$\eta(\mathbf{r}_n) = 4(2\pi)^{1/2} \cdot N_d \sum_{l_n, m_{l_n}} i^{l_n} \cdot Y_{l_n l_n}^{m_{l_n}}(\varrho_z) \cdot Y_{l_n}^{m_{l_n}*}(\varrho_z) \cdot G(|\mathbf{k}_p - \frac{1}{2}\mathbf{K}|) \cdot j_{l_n}(zr_n), \quad (58)$$

$$\mathbf{Z} = \mathbf{K} - \mathbf{k}_p, \quad (59)$$

$$G(|\mathbf{P}|) = \left[\frac{1}{P^2 + a^2} + \frac{1}{P^2 + (a+b)^2} \right]. \quad (60)$$

In order to take account of the spin states of all particles concerned, the spin-angle function of our system is taken to be, except for proton,

$$Y_{Jl_n s_j}^M = \sum_{m_{l_n}} \sum_{m_s} (l_n s m_{l_n} m_s | l_n s J M) Y_{l_n}^{m_{l_n}}(\varrho_n) \chi_{s m_s j},$$

where J and M are the total angular momentum and its z-component, s and m_s the channel spin wave function, which is given by

$$\chi_{s m_s j} = \sum_{m_j} \sum_{m_{s_n}} (j s_n m_j m_{s_n} | j s m_s) \chi_{j m_j} \cdot \chi_{s_n m_{s_n}}.$$

Here $\chi_{j m_j}$ and $\chi_{s_n m_{s_n}}$ are spin wave functions of the target nucleus and the neutron, respectively.

Then, Eq. (58) is modified as

$$\begin{aligned} \eta(\mathbf{r}_n) = 4(2\pi)^{1/2} N_d \sum K_{Jl_n m_{l_n} s_d s_p m_{s_p}}^M \cdot i^{l_n} \cdot \chi_{s_p m_{s_p}} \cdot \mathcal{Y}_{Jl_n s_j}^M \cdot Y_{l_n}^{m_{l_n}*}(\varrho_z) \\ \times G(|\mathbf{k}_p - \frac{1}{2}\mathbf{K}|) \cdot j_{l_n}(zr_n), \end{aligned} \quad (58b)$$

where the summation is extended with respect to J , M , m_{s_p} , l_n , m_{l_n} and the geometrical factor $K_{Jl_n m_{l_n} s_d s_p m_{s_p}}^M$ is given by

$$\begin{aligned} K_{Jl_n m_{l_n} s_d s_p m_{s_p}}^M = \sum_{\substack{m_{s_n} \\ m_s}} (s_n s_p m_{s_n} m_{s_p} | s_n s_p s_d m_{s_d}) (j s_n m_j m_{s_n} | j s m_s) \\ \times (l_n s m_{l_n} m_s | l_n s J M). \end{aligned} \quad (61)$$

Hence,

$$\begin{aligned} \omega_{\lambda n}^{II} = 4(2\pi)^{1/2} \cdot N_d \sum K_J^M \cdot i^{l_n} \chi_{s_p m_{s_p}} \cdot G(|\mathbf{k}_p - \frac{1}{2}\mathbf{K}|) \cdot Y_{l_n}^{m_{l_n}*}(\varrho_z) \\ \times (-b^2/2M) \int_{\Omega_{12}} \langle \Pi_{\lambda}^c | \chi_i(\tilde{z}) \mathcal{Y}_{Jl_n s_j}^M dS^{12} \left\{ \frac{\partial}{\partial r_n} j_{l_n}(zr_n) - \frac{\partial}{\partial r_n} [\log h_{l_n}^{(1)}(k_n r_n)] \cdot j_{l_n}(zr_n) \right\} \Big|_{r_n=r_0} \end{aligned} \quad (62)$$

by use of the boundary condition imposed on $|II_\lambda^e\rangle$,

$$(\partial/\partial r_n - \bar{f}) |II_\lambda^e\rangle = 0, \quad (r_n = r_0) \quad (63a)$$

$$\bar{f} = \left[\frac{d}{dr_n} \log b_{l_n}^{(2)}(k_n r_n) \right]_{r=r_0} \quad (63b)$$

and

$$\hbar^2 k_n^2 / 2M = E_n \quad \text{for the energy of the captured neutron.}$$

The (amplitude of) neutron width is defined as

$$\gamma_{\lambda s, l_n}^{II} = -(\hbar^2 k_n / M_n)^{1/2} \cdot \int_{s_{12}} \langle II_\lambda | \cdot \chi_i(\hat{z}) \cdot \gamma_{j_{l_n}, s_j}^M dS^{12} \left\{ \frac{\partial}{\partial r_n} j_{l_n}(k_n r_n) - f_l \cdot j_l(k_n r_n) \right\}_{r_n=r_0}$$

which, having a dimension of (energy)^{1/2}, is the u_{l_n} in Kapur and Peierls' theory in a generalized sense¹²⁾ in which the composition of angular momenta is performed in accordance with the concept of channel spin. Therefore, $\omega_{\lambda n}^{II}$ can be written in terms of the neutron width as follows:

$$\begin{aligned} \omega_{\lambda n}^{II} &= 2(2\pi)^{1/2} \cdot N_{\alpha} r_0^2 \sum K^J \cdot i^{l_n} \cdot \chi_{s_p m_{s_p}} \cdot G(|\mathbf{k}_p - \frac{1}{2}\mathbf{K}|) \cdot Y_{l_n}^*(Q_z) \\ &\times (\hbar^2 k_n / M_n)^{1/2} \cdot \gamma_{\lambda s, l_n}^J \left[\frac{\partial}{\partial r_n} j_{l_n}(z r_n) \cdot b_{l_n}^{(1)}(k_n r_n) - j_{l_n}(z r_n) \cdot \frac{\partial}{\partial r_n} b_{l_n}^{(1)}(k_n r_n) \right]_{r=r_0} \end{aligned} \quad (64)$$

(ii) Integration with respect to k_p .

Since the boundary condition of Kapur and Peierls are apparently Hermitian in the closed channel, W_λ^{II} will become real quantity E_λ^{II} and $N_\lambda^{II} = 1$ in the closed channels. Then, there appear poles of integral in the upper domain on the k_p -plane because of the small positive number ϵ . Then it is sufficient only to calculate their residues and to take a limit $\epsilon \rightarrow 0$.

(iii) Next we shall calculate the first factor in each of the sum of (52).

Since ϕ_f is the product of the wave functions of the residual nucleus and of the plane wave of the emitted proton $|k_p'\rangle$, we have

$$\begin{aligned} T_p &\equiv \langle \phi_f | L_p \left[1 + \frac{1}{E + i\epsilon - H_{II} - T_p - L_p} L_p \right] | k_p II_\lambda \rangle \\ &= \langle k_p' | L_p \left[1 + \frac{1}{\mathcal{E}_p' + i\epsilon - T_p - L_p} L_p \right] | k_p \rangle \langle \chi_f | II_\lambda \rangle, \end{aligned} \quad (65)$$

with $\mathcal{E}_p' = \hbar^2 k_p'^2 / 2M_p$.

Therefore, it has a finite value only if the state $|II_\lambda\rangle$ is identical with the residual nucleus χ_f and $k_p = k_p'$. Otherwise it becomes zero. Then, it follows, from (65),

$$T_p = \langle k_p' | L_p \left[1 + \frac{1}{\mathcal{E}_p' + i\epsilon - T_p - L_p} L_p \right] | k_p \rangle,$$

which represents evidently T-matrix of proton scattering by the potential L_p and can be

written in terms of the phase shift δ_{l_p} as follows.

$$T_p = 2\hbar^2 / M_p k_p \cdot (2\pi)^{1/2} \cdot \sum_{l_p, m_{l_p}} \sin \delta_{l_p} \cdot e^{i\delta_{l_p}} \cdot Y_{l_p}^{m_{l_p}}(\Omega_{k_p}) \cdot Y_{l_p}^{m_{l_p}*}(\Omega_{k_p}'), \quad (66)$$

where Ω_{k_p} and Ω_{k_p}' denote the angles of k_p and k_p' with respect to the direction of K .

By substituting (61) and (66) in (52) and, further, performing the integration over k_p , we get finally

$$\begin{aligned} T_{pd}^{(s')} &= 16\pi^2 \cdot N_d \cdot r_0^2 (\hbar^2 k_n / M_n)^{1/2} \cdot i^{l_n} \cdot T_{\lambda s, l_n}^{J, m_{l_n}} \sum_{l_p, m_{l_p}} K^J \cdot \sin \delta_{l_p} \cdot e^{i\delta_{l_p}} \cdot Y_{l_p}^{m_{l_p}*}(\Omega_{k_p}') \\ &\times \int d\Omega_{k_p} \cdot Y_{l_p}^{m_{l_p}}(\Omega_{k_p}) \cdot Y_{l_n}^{m_{l_n}*}(\Omega_z) G(|k_p - \frac{1}{2}K|) \left[\frac{\partial}{\partial r_n} j_{l_n}(zr_n) \cdot h_{l_n}^{(1)}(k_n r_n) \right. \\ &\quad \left. - j_{l_n}(zr_n) \frac{\partial}{\partial r_n} h_{l_n}^{(1)}(k_n r_n) \right]_{r=r_0}, \end{aligned} \quad (67)$$

where l_n is fixed to a certain value.

If $T_p^{(c)}$ is written in the form of pole expansion as (50), the potential L_p should be taken to be a hard sphere in which the effect of compound nucleus formation by the proton is entirely neglected. For the analysis of experimental results, it seems more reasonable for L_p to be taken as a complex potential which is determined by scattering experiments of a proton by the residual nucleus. In this case the process may be interpreted as follows; first the neutron is captured by the target nucleus to form residual nucleus and then the proton suffers the so-called shape elastic scattering⁷⁾ by the residual nucleus.

Corresponding to the direct interaction in nucleon (inelastic) scattering⁷⁹⁾, the indirect stripping process may be considered, that is, the intermediate states of the nucleus which is formed by the neutron capture are transformed to the residual nucleus by the direct excitation or by the direct de-excitation caused by the proton. The T-matrix element of this process can be evaluated through replacing the factor T_p in (52) by the T-matrix of the direct interaction. This means to pick up the appropriate terms, responsible for the mentioned process, in the T-matrix of the compound nucleus formation by deuteron derived in section 4.

§ 6. Final comments

In the preceding sections, it has been shown that the T-matrix of (d, p) reaction can be written as

$$T_{pd} = T_{pd}^{(s)} + T_{pd}^{(s')} + T_{pd}^{(c)} + \tilde{T}_{pd}^{(c)}$$

The cross section of this process, therefore, is obtained through

$$d\sigma(d, p)/d\Omega = (1/2\pi^2 \hbar^4) M_p M_d (k_p'/K) |T_{pd}|^2$$

and by taking the average with respect to the initial states and performing summation over the final states. If the compound nucleus excitation is so high that many levels

overlap, the interference term between the stripping and compound nucleus process will vanish provided signs of various γ are uncorrelated. This may be the reason why the observed angular distributions of (d, p) reaction actually show the behavior of Butler's characteristic curve superimposed by the characteristic curve of the statistical theory. At low excitation, however, the interference term will not vanish but play an important role in the (d, p) reaction.

So far, the ordinary deuteron channel has been neglected in our calculation. To take it into account exactly may hardly be possible in our formulation, because it does not use the orthogonality of channels explicitly. We shall discuss on its effect briefly. As is discussed in the beginning of Section 3, it is necessary to distinguish the deuteron channel wave function ψ_a^c from that of reaction product ψ_a^r . Then, the solution of the Schrödinger equation in the deuteron channel (in the external region) is given by

$$\psi_a^r [cF(KR) + c'G(KR)], \quad \psi_a^r = \psi_a^c + \phi$$

where ϕ is orthogonal to ψ_a^c on the surface S . In the preceding calculation we have assumed $\psi_a^c = 0$. If $\psi_a^c \neq 0$, ϕ becomes smaller, in magnitude, than in the case of $\psi_a^c = 0$ because of the normalization of flux. Thus the process $T_{pi}^{(s')}$ may be slightly reduced. Since the separation $\psi_a^c + \phi$ depends, of course, on the internal coordinate of deuteron, it may be supposed that its angular distribution is also modified. However, it seems plausible to assume that ψ_a^c is small compared with ϕ by taking account of rather small region of deuteron channel in configuration space.

The extension of this theory to the (d, np) reaction and the (elastic and inelastic) scattering of deuteron can easily be performed, which results will appear soon in this journal.

The author wishes to express his sincere thanks to Professor T. Muto for his encouragement and valuable comments and for the careful reading of the manuscript before publication. His thanks are also due to M. Kawai and M. Nagasaki for their stimulating discussions.

References

- 1) S. T. Butler, Proc. Roy. Soc. **A208** (1951), 559.
- 2) Bhatia, Huang, Huby and Newns, Phil. Mag. **43** (1952), 485.
R. Huby, Proc. Roy. Soc. **A215** (1952), 385.
S. Yoshida, Prog. Theor. Phys. **10** (1953), 1.
E. Gerjouy, Phys. Rev. **91** (1953), 645.
- 3) W. Tobocman, Phys. Rev. **94** (1954), 1655.
- 4) Stratton, Blair, Famularo and Stuart, Phys. Rev. **98** (1955), 629.
- 5) Cohen and Folk, Phys. Rev. **84** (1951), 173.
- 6) H. A. Bethe and S. T. Butler, Phys. Rev. **85** (1952), 1045.
S. Okai and M. Sano, Prog. Theor. Phys. **14** (1955), 399.
- 7) K. G. Thomas, Phys. Rev. **100** (1955), 45.

- 8) T. Teichman and E. P. Wigner, Phys. Rev. **87** (1952), 123.
- 9) P. L. Kapur and R. Peierls, Proc. Roy. Soc. **A166** (1938), 277.
- 10) B. A. Lippman and J. Schwinger, Phys. Rev. **72** (1950), 469.
- 11) E. P. Wigner and L. Eisenbud, Phys. Rev. **72** (1950), 29.
- 12) M. Kawai and M. Nagasaki, to be published.
- 13) S. Hayakawa, Proc. International Conference on Theoretical Physics, Kyoto, Japan 1953.
- 14) M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91** (1953), 398.
Fujimoto, Hayakawa and Nishijima, Prog. Theor. Phys. **10** (1953), 113.
- 15) P. M. Morse and H. Feshbach, *Method of Theoretical Physics* (Technology Press, Massachusetts 1946 Chap. 7).
- 16) N. C. Francis and K. M. Watson, Phys. Rev. **93** (1954), 313.
- 17) Feshbach, Porter and Weisskopf, Phys. Rev. **96** (1954), 448.
- 18) H. Ui, Prog. Theor. Phys. **14** (1955), 75.
- 19) Austern, Butler and McManus, Phys. Rev. **92** (1954), 350.
S. Hayakawa, Trans. Bose Res. Inst. **20** (1955), 21.
S. Hayakawa and S. Yoshida, Prog. Theor. Phys. **14** (1955), 1.

Nuclear Deformability and Shell Structure

Toshio MARUMORI, Shôta SUEKANE*†
and Atsuko YAMAMOTO**

Research Institute for Fundamental Physics, Kyoto University, Kyoto

*Department of Physics, Kyoto University, Kyoto**

*Institute of Theoretical Physics, Nagoya University, Nagoya***

(Received June 25, 1956)

The connection between the surface rigidity of the nuclear core, which has been accounted for in terms of the surface tension of liquid drop nuclei, and its proper shell structure is discussed by using the method of the quantum mechanical description of the collective motion, which has been proposed by one of the present authors and others (§ 2). On the basis of such a consideration the surface rigidity of the core can be calculated, provided that the shell model is valid for the behaviour of particles forming the core (§ 3 and § 4). The noticeable features of our results obtained are that the calculated values of rigidity of cores are, in general, considerably larger than those due to the hydrodynamical estimation, and are closely related to the proper shell structures of cores (§ 4). Such a characteristic variation of the rigidity of cores is discussed in detail by comparing with the quadrupole moments of the "core \pm one extra-particle type" nuclei. Theoretical quadrupole moments thus obtained finely explain the observed values which have been noticed to depend on the shell structure (§ 5). Some discussions are devoted to the rigidity of a core with large deformation (§ 6).

§ 1. Introduction

The striking success of the nuclear collective model^{1), 2)} in accounting for many features of the nuclear structure has stimulated one's interest in some of underlying problems. One of these is the study of the relation between the surface rigidity of cores and the proper shell structures of them.

A simple hydrodynamical model^{1), 3), 4), 5), 6)} which replaces the core by a phenomenological liquid drop gives deformations considerably larger, sometimes by one order of magnitude, than empirical values deduced from nuclear quadrupole moments. Still more critical for this model is that it can not explain the rapid variation of the rigidity of cores, which must be accepted as an empirical fact. In particular, it should be noted⁷⁾ that this model is completely helpless in explaining the great variation of the surface rigidity in different isotopes of the same element, which can be seen from the quadrupole moment ratios of isotopes with the same nuclear spin.

Such circumstances make clear the phenomenological aspect of the collective model based on the simple hydrodynamical assumption. In fact, although we know that the strong coupling scheme of the collective model is successful for mass numbers between 155 and 185 and above 225, and that the weak coupling scheme applies best to the nuclei

†) Now at Department of Physics, Osaka City University.

in the proximity of major closed shells, we can not explain in the model itself why such treatments can be successful for those nuclei. This is due to the fact that in the collective model such treatments are determined by the strength of interaction between the core and extra-particles, which depends essentially on the rigidity of the core.

On the other hand, estimates based on the simple model^{(8),(9),(10)} in which particles are regarded as moving independently in a deformed potential have given much smaller deformations. However, such a deformed potential model has not been discussed in connection with Bohr-Mottelson's (B-M) formulation⁽¹⁾ on the collective model.

Recently, the method of the quantum mechanical description of the collective motion which gives the foundation of the nuclear collective model has been proposed^{(11),(12),(13)}, and it has been made clear⁽¹⁴⁾ that the difference between B-M's and Hill-Wheeler's descriptions of the nuclear collective model is nothing but that of the representation in the description of the nuclear states except for several trivial points.

By using this method, therefore, we can faithfully follow B-M's formulation of the collective model and can make clear the connection between the surface rigidity of a core and the proper shell structure of the corresponding core (§ 3). Then, we can find that in B-M's model the adiabatic approximation is used for the particles forming a core, and that the surface rigidity of a core is given by the second order coefficient of expansion of the energy eigen-value of particles forming the core, with respect to the deformation parameters of the core.

Provided that the shell model assumption is valid for the behaviour of such particles, therefore, the method of the deformed potential model can be applied to the calculation of surface rigidity of the corresponding core. Using Moszkowski's technique⁽¹⁰⁾, the calculation of the rigidity is made in § 3 and § 4.

The striking features of our results obtained are that the calculated values of rigidity of cores are, in general, considerably larger than those of the hydrodynamical estimation, and that there is a characteristic dependence of the rigidity upon the proper shell structures of cores (Fig. III). And this feature of rigidity can explain qualitatively why the strong coupling or weak coupling scheme of the collective model is successful with specific nuclei (§ 4).

More detailed discussions on such a characteristic variation of the rigidity of cores are done in § 5 by comparing with the quadrupole moments of "core \pm one extra-particle type" nuclei. Thus we can find that such a variation plays an essential part in the variation of the quadrupole moments, as shown in Table II. In particular, it should be emphasized that it can explain the great variations of the surface rigidity in the isotopes; S^{33} (neutrons: 16-shell+1) and S^{35} (neutrons: 20-shell-1), and Ga^{69} (neutrons: 38-shell) and Ga^{71} (neutrons: 40-shell). The differences between the quadrupole moments of pairs of these isotopes have never been explained in terms of the hydrodynamical model or of the method of configuration mixing.⁽¹⁵⁾

Some discussions are devoted in § 6 to the rigidity of core in the case of the large deformation.

§ 2. Formulation of the collective model

In order to clarify the connection between the surface rigidity of cores and the proper shell structures of them, it is necessary first of all to establish the formulation of the collective model. On this point, the method of the quantum mechanical description of the collective motion which has been formulated by one of the present authors^{(11), (14)} and others^{(12), (13)} is of use. Keeping close contact with B-M's description of the collective model, we shall develop this method in this section.

The Schrödinger equation which describes a nuclear system is

$$H\Phi \equiv \left[\sum_{i=1}^A \mathbf{p}_i^2 / 2M + V(\mathbf{x}_1, \dots, \mathbf{x}_A) \right] \Phi(\mathbf{x}_1, \dots, \mathbf{x}_A) = E\Phi(\mathbf{x}_1, \dots, \mathbf{x}_A), \quad (2.1)$$

where M is the nucleon mass and V the interaction potential between nucleons. Now we introduce the collective coordinates $\alpha_{lm}^{(*)}$ and start from the following equation,

$$\left[\sum_{i=1}^A \mathbf{p}_i^2 / 2M + V(\mathbf{x}_1, \dots, \mathbf{x}_A) \right] \Phi'(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) = E\Phi'(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha), \quad (2.2)$$

which is equivalent to (2.1) under the condition,

$$\alpha_{lm} \Phi'(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) = 0. \quad (2.2')$$

We transform (2.2) and (2.2') into the "collective representation" by the unitary transformation,

$$\begin{aligned} U &= \exp[i(M/\hbar) \int \rho(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}] \\ &= \exp[i(M/\hbar) \int \sum_{i=1}^A \delta(\mathbf{x} - \mathbf{x}_i) \phi(\mathbf{x}) d\mathbf{x}] \\ &= \exp[i(M/\hbar) \sum_{i=1}^A \phi(\mathbf{x}_i)]. \end{aligned} \quad (2.3) **$$

Here $\phi(\mathbf{x})$ is the velocity potential of an irrotational and incompressible carrier fluid and is represented by

$$\phi(\mathbf{x}) = \sum_{lm} \beta_{lm} r^l Y_{lm}(\theta, \varphi).$$

As is well known in the liquid drop model of the nucleus, β_{lm} are related to the canonical conjugate quantities π_{lm} of the coordinates α_{lm} through the relation,

$$\begin{aligned} \pi_{lm} &= B_l \cdot l \cdot R_0^{l-2} \cdot \beta_{lm}^*, \\ B_l &= l^{-1} \cdot (3/4\pi) \cdot AMR_0^2. \end{aligned}$$

*) α_{lm} are the expansion parameters of the nuclear surface defined by

$$R(\theta, \varphi) = R_0 [1 + \sum_{lm} \alpha_{lm} Y_{lm}(\theta, \varphi)]$$

where R_0 is the equilibrium radius.

**) According to Eden and Francis⁽¹⁶⁾ this corresponds to the "model operator" into the collective model. Here note that $\sum_{i=1}^A \phi(\mathbf{x}_i)$ is a symmetric function of \mathbf{x}_i , so that the "individual particles" in the "collective representation" obey the Fermi statistics in the same manner as nucleons.

In this "collective representation", (2.2) and (2.2') are expressed as

$$\begin{aligned} & \left[\sum_{i=1}^A \mathbf{p}_i^2 / 2M + (M/2) \cdot \int \rho(\mathbf{x}) (\text{grad } \phi(\mathbf{x}))^2 d\mathbf{x} + V(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) \right. \\ & \quad \left. + M \int \mathbf{v}(\mathbf{x}) \cdot \text{grad } \phi(\mathbf{x}) d\mathbf{x} \right] \Psi(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) = E \Psi(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha), \quad (2.4)^{*} \\ & \quad [\alpha_{lm} - \sum_{i=1}^A F_{lm}(\mathbf{x}_i)] \Psi(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) = 0, \quad (2.4') \end{aligned}$$

where

$$\sum_{i=1}^A F_{lm}(\mathbf{x}_i) = (4\pi/3A) \sum_{i=1}^A (r_i/R_0)^l Y_{lm}^*(\theta_i, \varphi_i)^{**}$$

and

$$\mathbf{v}(\mathbf{x}) = (1/2M) \sum_{i=1}^A [\mathbf{p}_i \delta(\mathbf{x} - \mathbf{x}_i) + \delta(\mathbf{x} - \mathbf{x}_i) \mathbf{p}_i],$$

which is the velocity operator of the individual particles.

Eq. (2.4) with (2.4') is equivalent to B-M's except for several trivial points as discussed in the following. We can easily ascertain that this representation is used in B-M's description, by calculating any physical operator in this representation.***)

1) Collective oscillations of particles forming a shell structure

The relationship of the motion of individual particles and the collective motion, which is described by eq. (2.4) with (2.4'), is especially simple if the frequencies for particle excitation are large compared with the frequencies of the collective type of motion. In this case, the adiabatic approximation can be used and the wave function of (2.4) is represented approximately by

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha) = \phi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) \chi(\alpha), \quad (2.5)$$

where $\chi(\alpha)$ describes the oscillation of a nucleus as a whole. Here $\phi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha)$ is the solution of

$$\left[\sum_{i=1}^A \mathbf{p}_i^2 / 2M + V(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) \right] \phi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) = E^{\text{in}}(\alpha) \phi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha)$$

and must satisfy approximately

$$\begin{aligned} & \int \phi^*(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) \rho(\mathbf{x}) \phi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) d\mathbf{x}_1 \dots d\mathbf{x}_A \\ & = \begin{cases} \rho_0 (= 3A/4\pi R_0^3) & ; \text{ inside the deformed nucleus} \\ 0 & ; \text{ outside the deformed nucleus} \end{cases} \quad (2.6) \end{aligned}$$

*) Strictly speaking, the α -dependence of V is not uniquely defined by this transformation, but this dependence can be approximately given by Tomonaga's procedure¹²⁾ or Miyazima's transformation.¹³⁾

**) This function is just the expression for the collective parameters in terms of the individual particles. (cf. 1), p. 10).

***) It is clear that Tolhoek's¹⁷⁾ and Coester's¹⁸⁾ analyses of B-M's model are equivalent to solving (2.4) with (2.4') by use of some assumptions; if we assume that $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_A, \alpha)$ has a form, $\delta(\alpha - \Sigma F(\mathbf{x}_i)) \phi(\mathbf{x}_1, \dots, \mathbf{x}_A) g(\alpha)$ and calculate an equation which $g(\alpha)$ satisfies, we obtain the same results as Tolhoek's and Coester's.

because of the condition (2.4'). (2.6) ensures

$$\int \psi^*(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) \sum_{i=1}^A F_{lm}(\mathbf{x}_i) \psi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) d\mathbf{x}_1 \dots d\mathbf{x}_A = \alpha_{lm}$$

so that the subsidiary condition for $\chi(\alpha)$:

$$[\alpha_{lm} - \int \psi^*(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) \sum_{i=1}^A F_{lm}(\mathbf{x}_i) \psi(\mathbf{x}_1, \dots, \mathbf{x}_A; \alpha) d\mathbf{x}_1 \dots d\mathbf{x}_A] \chi(\alpha) = 0$$

becomes an identical equation. Therefore the collective motion is described by only one equation,

$$[(M/2) \int \langle \rho(\mathbf{x}) \rangle (\text{grad } \phi(\mathbf{x}))^2 d\mathbf{x} + E''(\alpha)] \chi(\alpha) = E \chi(\alpha). \quad (2.7)^*$$

If the particle structure prefers spherical symmetry, $E''(\alpha)$ may be expanded around the equilibrium state ($\alpha_{lm} = 0$):

$$E''(\alpha) = E''^{(0)} + \sum_{lm} E''^{(2)}_{lm} |\alpha_{lm}|^2.$$

Then (2.7) is reduced to

$$[\sum_{lm} (1/2B_l) |\pi_{lm}|^2 + \sum_{lm} E''^{(2)}_{lm} |\alpha_{lm}|^2] \chi(\alpha) = (E - E''^{(0)}) \chi(\alpha). \quad (2.7')$$

This equation is nothing but that for the surface motion given by B-M. (cf. 1), II. 5)

2) Coupling to particle motion

If the particles possess modes of excitation with frequencies smaller than or comparable with the collective frequencies, the simple separation of collective and particle motion such as (2.5) is no longer possible. Since the particle structure can be non-adiabatically excited in this case, the nucleus must be described in terms of a coupled system of collective and particle degrees of freedom.

In the treatment of this case, B-M have adopted the picture of "core + extra-particles." Namely, the degrees of freedom of the extra-particles represent the low frequency modes of excitation of the particle structure, associated with the particles in the last unfilled levels. And the particles forming a closed shell (core) manifest themselves only through the collective motion of the core.

Conforming to this picture, our description is reformulated as follows. We perform the unitary transformation (2.3) only for the quantities which are concerned with the particles constituting the closed shell, so that the Schrödinger equation in the "collective representation" is given by

$$\begin{aligned} & \left[\sum_{i=1}^N \mathbf{p}_i^2 / 2M + (M/2) \int \rho^{\text{core}}(\mathbf{x}) (\text{grad } \phi(\mathbf{x}))^2 d\mathbf{x} + V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha) \right. \\ & \left. + M \int \mathbf{v}^{\text{core}}(\mathbf{x}) \cdot \text{grad } \phi(\mathbf{x}) d\mathbf{x} + \sum_{j=N+1}^A \mathbf{p}_j^2 / 2M + V^{(2)}(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A) \right] \end{aligned}$$

*) The term including $\mathbf{v}(\mathbf{x})$ in (2.4) vanishes, when we take its expectation value in ψ .

$$+V^{(3)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha)]\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) \quad (2.8)$$

$$=E\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha),$$

$$[\alpha_{lm} - \sum_{i=1}^N F_{lm}(\mathbf{x}_i)]\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) = 0, \quad (2.8')$$

where $\mathbf{x}_1, \dots, \mathbf{x}_N$ denote the coordinates of the particles which constitute the closed shell (core) and $\mathbf{x}_{N+1}, \dots, \mathbf{x}_A$ denote those of the particles in the last unfilled levels (extra-particles). Prior to the unitary transformation, we divided the interaction potential between the particles into three parts. The first part is the interaction potential between the particles forming the core, and the second is that between the extra-particles. The third is the interaction potential between the particles forming the core and the extra-particles. These in the "collective representation" are denoted by $V^{(1)}$, $V^{(2)}$, and $V^{(3)}$ respectively.*

In the present case, the adiabatic approximation is possible only for the core part. Therefore, the approximate wave function which satisfies eq. (2.8) with (2.8') may be of the form:

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) = \psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha) \chi(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha). \quad (2.9)$$

$\psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha)$ is the solution of the equation,

$$\left[\sum_{i=1}^N \mathbf{p}_i^2 / 2M + V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha) \right] \psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha) = E^{core}(\alpha) \psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha), \quad (2.10)$$

and must satisfy approximately in the same manner as (2.6)

$$\begin{aligned} & \int \psi^{core*}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha) \rho(\mathbf{x}) \psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha) d\mathbf{x}_1 \dots d\mathbf{x}_N \\ &= \begin{cases} \rho_0 (=3N/4\pi R_0^3) & ; \text{ inside the deformed core} \\ 0 & ; \text{ outside the deformed core} \end{cases} \end{aligned} \quad (2.11)$$

because of the condition (2.8'). In this case $\chi(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A; \alpha)$ is no longer subject to the subsidiary condition and satisfies the equation;

$$\begin{aligned} & [H_{surf}(\alpha) + H_p(\mathbf{x}_p) + H_{int}(\mathbf{x}_p, \alpha)] \chi(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) \\ &= (E - E_0^{core}) \chi(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha), \end{aligned} \quad (2.12)$$

where

$$\begin{aligned} H_{surf}(\alpha) &= (M/2) \int \langle \rho(\mathbf{x}) \rangle^{core} (\text{grad } \phi(\mathbf{x}))^2 d\mathbf{x} + (E^{core}(\alpha) - E_0^{core}), \\ H_p(\mathbf{x}_p) &= \sum_{j=N+1}^A \mathbf{p}_j^2 / 2M + \langle V^{(3)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) \rangle^{core} \Big|_{\alpha=0} \\ &\quad + V^{(2)}(\mathbf{x}_{N+1}, \dots, \mathbf{x}_A), \end{aligned} \quad (2.13)$$

$$H_{int}(\mathbf{x}_p, \alpha) = \frac{\partial}{\partial \alpha} \langle V^{(3)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \mathbf{x}_{N+1}, \dots, \mathbf{x}_A, \alpha) \rangle^{core} \Big|_{\alpha=0} \alpha.$$

*) The α -dependences of $V^{(1)}$ and $V^{(3)}$ can also be given approximately in the same way as that of V in (2.4), but we shall not touch this point here.

(2.13) is nothing but the Hamiltonian which was adopted by B-M as the starting point of the collective model, and eq. (2.12) is just the Schrödinger equation describing the collective model.

3) Surface rigidity of the core and its proper shell structure

From the above formulation of the collective model, we can easily find the relationship between the surface rigidity of the core and the proper shell structure of it. The eq. (2.10) describes the particle structure of the core, provided that the wave function $\psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha)$ satisfies (2.11) approximately. And the surface rigidity of the core is given by the second order coefficient of expansion of $E^{core}(\alpha)$ with respect to α . The problem to obtain the surface rigidity of the core is thus focussed to calculate $E^{core}(\alpha)$.

In the calculation of $E^{core}(\alpha)$, the specific type and form of the interaction potential, $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$, come at first into question. As a rule, a concrete form of $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$ can be obtained approximately by Tomonaga's procedure¹²⁾ or Miyazima's transformation¹³⁾ when the interaction potential between nucleons is given. In the present stage of the study of nuclear forces in a many nucleon system, however, it may not be so much significant to introduce the concrete form of $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$. It seems rather natural to assume the properties of $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$ through the model which gives an important guide in the interpretation of nuclear phenomena. The striking success of the nuclear shell model suggests that the particles may be considered as occupying states of binding characteristic to the independent particle motion in an averaged potential. Therefore, we assume here that $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$ has the following properties:

- i) In the zero order approximation, $V^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_N, \alpha)$ can be replaced by an averaged potential with the deformation described by α .*)
- ii) The form of this averaged potential is assumed to be of the square well, because the wave function $\psi^{core}(\mathbf{x}_1, \dots, \mathbf{x}_N; \alpha)$ must satisfy the condition (2.11) approximately.
- iii) Moreover, if necessary, we can add the amendments such as the $l \cdot s$ -coupling term and others to it. These amendments to the deformed averaged potential may be considered as a residual part of $V^{(1)}$.

§ 3. Method of calculations

Since we are calling the low energy states in the vicinity of the ground state into question, we shall adopt only $\alpha_{l=2}$ of $l=2$ for the deformation of a core. Moreover, since the observed rotational spectra indicate that the many nuclei prefer the cylindrical symmetry,

*) The result calculated by use of such a deformed averaged potential indicates that the first order expansion coefficient of $E^{core}(\alpha)$ becomes zero. This means that the magnitude of the deformation of the core without extra-particles:

$$(\chi^{core}(\alpha), \alpha \chi^{core}(\alpha)) = -E^{(1)core}/2E^{(2)core}$$

($E^{(1)core}$ and $E^{(2)core}$ are the first and second order expansion coefficients of $E^{core}(\alpha)$ respectively.) is zero. Therefore, B-M's assumption that the core prefers the spherical symmetry is very natural.

we shall perform the calculation by using a spheroidally deformed, averaged potential.

In this case, the wave function in (2.10) is represented by the Slater determinant of single particle wave functions in the spheroidal square well potential, and the energy eigen-value, $E^{\text{core}}(\alpha)$, in (2.10) is expressed by the sum of energy eigen-values, $E^s(\alpha)$, with respect to the single particle wave functions:

$$E^{\text{core}}(\alpha) = \sum_s E^s(\alpha). \quad (3.1)$$

Therefore, our problem to obtain $E^{\text{core}}(\alpha)$ is reduced in calculating the energy eigen-values of the Schrödinger equation with respect to the single particle in the spheroidal square well potential.

1) Moszkowski's transformation¹⁰⁾

The good quantum numbers of particle states in a spheroidal well are the parity and the component of angular momentum about the axis of deformation. Hereafter we shall denote this symmetry axis of deformation as z and the component of angular momentum about it as m .

According to Moszkowski's method, we shall calculate the energy eigen-values of a single particle in the spheroidal square well. First, the coordinate transformation:

$$x = D^{-1/2} x', \quad y = D^{-1/2} y' \quad \text{and} \quad z = Dz' \quad (3.2)^*)$$

is carried out. Here D is defined for the small deformation by

$$D = 1 + d, \quad d = \sqrt{5/4\pi} \beta, \quad (3.3)$$

where

$$\beta^2 = \sum_m |\alpha_{2m}|^2.$$

In this new coordinate system, the spheroidal square well potential becomes

$$V(r') = \begin{cases} 0 & \text{for } r' > R_0 \\ -V_0 & \text{for } r' < R_0 \end{cases}$$

which means the square well potential. And the Schrödinger equation of a single particle in this coordinate system becomes

$$(1/2M) (DP_{x'}^2 + DP_{y'}^2 + D^{-2} P_{z'}^2) \psi(x') + V(r') \psi(x') = E^s \psi(x'). \quad (3.4)$$

Because we are concerned with the surface rigidity of the core (closed shell) which prefers the spherical symmetry, it is sufficient to obtain $E^s(\alpha)$ to the second order with respect to d by using, as the zero order functions, the functions $\psi_{s0}^{(0)}(x')$ for the Hamiltonian (3.4) when $D=1$. To avoid confusion, we shall hereafter drop the "prime" of the newly transformed coordinates x' .

* This method has the same meaning as Miyazima's transformation in a sense.

2) *Determinations of the level scheme and the property of potential*

In our problem, the ordering and the distances of the single particle energy levels in the zero order approximation are of decisive importance. Therefore, we shall add the $\mathbf{l} \cdot \mathbf{s}$ -coupling^{*)} and \mathbf{l}^2 ^{**)} terms to $V(r)$ so that the level scheme for the zero order approximation reproduces approximately Klinkenberg's one¹⁾ which is based on the shell model. Thus (3.4) is rewritten as follows:

$$(H^0 + H')\psi(\mathbf{x}) = E^0\psi(\mathbf{x}), \quad (3.5)$$

$$\begin{aligned} H^0 &= \mathbf{p}^2/2M + V(r) - (\hbar^2/2MR_0^2)k\mathbf{l} \cdot \mathbf{s} + (\hbar^2/2MR_0^2)g\mathbf{l}^2, \\ H' &= (\mathbf{p}^2/2M)d^2 - 2(d-d^2)T_{20}(\mathbf{p}), \\ T_{20}(\mathbf{p}) &= (1/2M) \cdot (3p_z^2 - \mathbf{p}^2)/2, \quad H' = dH'^{(1)} + d^2H'^{(2)}, \\ dH'^{(1)} &= -2dT_{20}(\mathbf{p}), \quad d^2H'^{(2)} = d^2(\mathbf{p}^2/2M + 2T_{20}(\mathbf{p})). \end{aligned} \quad (3.6)$$

The \mathbf{l}^2 -term may mean a correction to the square well potential $V(r)$. Namely, this pushes up the states of higher angular momenta, so that it may be considered to play a role of interpolation to the oscillator potential. In this sense it seems natural to set up this term in the transformed system. The constant g of \mathbf{l}^2 -term and the constant k of $\mathbf{l} \cdot \mathbf{s}$ -coupling term are treated as parameters which are adjusted suitably so as to reproduce Klinkenberg's level scheme. The values of g and k are, therefore, chosen as dependent on energy levels, as shown in Fig. 1.^{***)}

For the sake of simplifying calculations, we shall hereafter replace $V(r)$ with an infinite square well and the effects of the finite wall height will be taken into account later.

3) *Expression of E^0*

The eigen-functions for H^0 , which are used as the zero order functions, are

$$\begin{aligned} \psi_{nljm}^0 &= c j_n(\omega_{nl}, r/R_0) \left[\sqrt{(j+m)/2j} \alpha Y_{j-1/2}^{m-1/2} + \sqrt{(j-m)/2j} \beta Y_{j-1/2}^{m+1/2} \right] \quad \text{for } j=l+1/2 \\ \psi_{nljm}^0 &= c j_n(\omega_{nl}, r/R_0) \left[\sqrt{(j-m+1)/(2j+2)} \alpha Y_{j+1/2}^{m-1/2} - \sqrt{(j+m+1)/(2j+2)} \beta Y_{j+1/2}^{m+1/2} \right] \\ &\quad \text{for } j=l-1/2, \end{aligned}$$

in which c is a normalization constant, α and β are Pauli spin functions, and $j_n(\omega_{nl}, r/R_0)$ is the usual spherical Bessel function. And the eigen-values of H^0 are given by

*) Whether or not the adoption of the $\mathbf{l} \cdot \mathbf{s}$ -coupling term in such a transformed coordinate system is legitimate may depend on its origin. Since the $\mathbf{l} \cdot \mathbf{s}$ -coupling term used in the shell model is set up under the assumption of the "spherically symmetric" averaged potential, the adoption of $\mathbf{l} \cdot \mathbf{s}$ -coupling term in such a transformed coordinate system would not always be nonsense.

**) After the work was finished, it has come to our notice that S. G. Nilsson [Kgl. Danske Viedenskab Selskab, Mat.-fys. Medd 29, No. 16 (1955)] used this \mathbf{l}^2 term for the same purpose as ours. The authors are indebted to Dr. T. Tamura for his kind information of this point.

***) According to the physical meaning of \mathbf{l}^2 term, we adopted here the same values of g for the degenerated states in an oscillator potential.

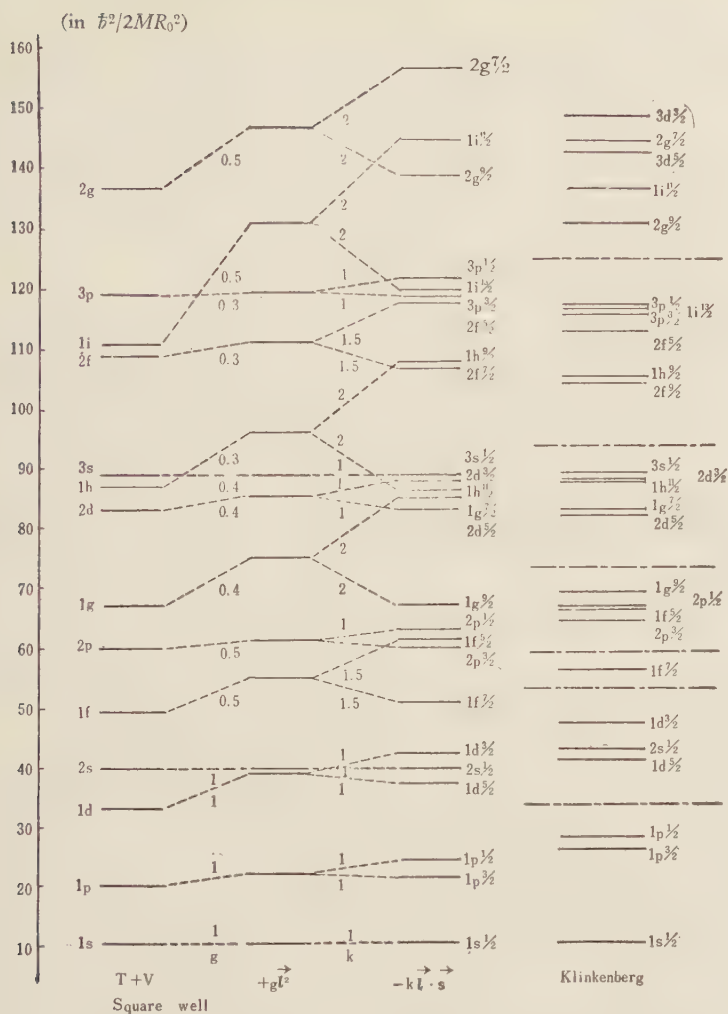


Fig. 1 Level scheme.

$$W_{nlj}^0 = \begin{cases} (\hbar^2/2MR_0^2) \omega_{nl}^2 + (\hbar^2/2MR_0^2) gl(l+1) + (\hbar^2/2MR_0^2) k(l+1) & \text{for } j=l-1/2 \\ (\hbar^2/2MR_0^2) \omega_{nl}^2 + (\hbar^2/2MR_0^2) gl(l+1) - (\hbar^2/2MR_0^2) kl & \text{for } j=l+1/2 \end{cases} \quad (3.7)$$

where ω_{nl} denotes the n -th root of the spherical Bessel function of order l . Therefore, E_{nljm}^s which is calculated to the second order with respect to d by using H' as a perturbation is given as follows:

$$\begin{aligned} E_{nljm}^s &= W_{nlj}^0 + E_{nljm}^{(1)} d + E_{nljm}^{(2)} d^2, \\ E_{nljm}^{(1)} &= \langle n l j m | H'^{(1)} | n l j m \rangle = -2 \langle n l j m | T_{20}(\mathbf{p}) | n l j m \rangle \\ &= -2 (\hbar^2/2MR_0^2) \omega_{nl}^2 \langle l j m | P_2(\theta) | l j m \rangle, \end{aligned} \quad (3.8)$$

$$\begin{aligned}
 E_{nljm}^{(2)} &= \langle n l j m | H'^{(2)} | n l j m \rangle + \sum_k |\langle n l j m | H'^{(1)} | n_k l_k j_k m_k \rangle|^2 / (W_{nlj}^0 - W_k^0) \\
 &= (\hbar^2/2MR_0^2) \omega_{nl}^2 [1 + 2 \langle l j m | P_2(\theta) | l j m \rangle] \\
 &\quad + \sum_k |\langle n l j m | H'^{(1)} | n_k l_k j_k m_k \rangle|^2 / (W_{nlj}^0 - W_k^0). \quad (3.9)
 \end{aligned}$$

The second term in the right hand side of (3.9) can be evaluated in reference to the well-known properties of spherical Bessel functions:

$$\begin{aligned}
 &\sum_k |\langle n l j m | H'^{(1)} | n_k l_k j_k m_k \rangle|^2 / (W_{nlj}^0 - W_k^0) \\
 &= 4 (\hbar^2/2MR_0^2) \sum_{n'} \left[\frac{4 (2l+3)^2 \omega_{nl}^2 \omega_{n'l, l+2}^2}{(\omega_{nl}^2 - \omega_{n'l, l+2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l+2, j+2, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l+2, j+2}^0) / (\hbar^2/2MR_0^2)} \right. \\
 &\quad + \frac{4 (2l+3)^2 \omega_{nl}^2 \omega_{n'l, l+2}^2}{(\omega_{nl}^2 - \omega_{n'l, l+2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l+2, j+1, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l+2, j+1}^0) / (\hbar^2/2MR_0^2)} \\
 &\quad + \omega_{nl}^4 \delta_{nn'} \frac{|\langle l j m | P_2(\theta) | l j - 1, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l-1}^0) / (\hbar^2/2MR_0^2)} \\
 &\quad \left. + \frac{4 (2l-1)^2 \omega_{nl}^2 \omega_{n'l, l-2}^2}{(\omega_{nl}^2 - \omega_{n'l, l-2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l-2, j-2, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l-2, j-2}^0) / (\hbar^2/2MR_0^2)} \right] \quad (3.10) \\
 &\quad \text{for } j=l+1/2,
 \end{aligned}$$

$$\begin{aligned}
 &\sum_k |\langle n l j m | H'^{(1)} | n_k l_k j_k m_k \rangle|^2 / (W_{nlj}^0 - W_k^0) \\
 &= 4 (\hbar^2/2MR_0^2) \sum_{n'} \left[\frac{4 (2l+3)^2 \omega_{nl}^2 \omega_{n'l, l+2}^2}{(\omega_{nl}^2 - \omega_{n'l, l+2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l+2, j+2, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l+2, j+2}^0) / (\hbar^2/2MR_0^2)} \right. \\
 &\quad + \omega_{nl}^4 \delta_{nn'} \frac{|\langle l j m | P_2(\theta) | l, j+1, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, j+1}^0) / (\hbar^2/2MR_0^2)} \\
 &\quad + \frac{4 (2l-1)^2 \omega_{nl}^2 \omega_{n'l, l-2}^2}{(\omega_{nl}^2 - \omega_{n'l, l-2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l-2, j-1, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l-2, j-1}^0) / (\hbar^2/2MR_0^2)} \\
 &\quad \left. + \frac{4 (2l-1)^2 \omega_{nl}^2 \omega_{n'l, l-2}^2}{(\omega_{nl}^2 - \omega_{n'l, l-2}^2)^2} \frac{|\langle l j m | P_2(\theta) | l-2, j-2, m \rangle|^2}{(W_{nlj}^0 - W_{n'l, l-2, j-2}^0) / (\hbar^2/2MR_0^2)} \right] \\
 &\quad \text{for } j=l-1/2. \quad (3.11)
 \end{aligned}$$

4) Approximations in numerical calculations

From (3.8), (3.9), (3.10) and (3.11), the following facts are easily understood:

- i) The coupling between $\psi_{nljm}^{(0)}$ and $\psi_{n'l'j'm'}^{(0)}$ by a spheroidal distortion can arise only when $m-m'=0$ and $l-l'=0$ or ± 2 .
- ii) The matrix elements of $H'^{(1)}$ between the states having the different values of n vanish if the orbital angular momenta are the same.
- iii) The matrix elements $|\langle n l j m | H'^{(1)} | n' l' j' m \rangle|$, in which $l-l'=\pm 2$ and $|n-n'| \geq 2$, are very small because of the large energy denominator $|\omega_{nl}^2 - \omega_{n'l'}^2|$. Therefore, the main contributions to $E_{nljm}^{(2)}$ result from the coupling of $\psi_{nljm}^{(0)}$ with states $\psi_{n'l'j'm}^{(0)}$ in which $l'=l$ or $l \pm 2$ and $n'=n$ or $n \pm 1$.

iv) Of these main contributions to $E_{nljm}^{(2)}$, the contributions resulting from the coupling with the states in the proximity of the n, l, j -level are especially large. And the values are strongly dependent on the level spacing $\Delta_{nlj-n'l'j'}$.

According to the above considerations, we shall adopt the following approximation and treatment in calculating the surface rigidity of a neutron or proton core (closed shell) under consideration,

$$C = \partial^2 E^{core}(\alpha) / \partial \alpha^2, \alpha=0 = (5/2\pi) \cdot \sum_{nljm}^{closed\ shell} E_{nljm}^{(2)} \quad (3 \cdot 12)$$

(a) The sums over n' in (3·10) and (3·11) are performed only for $n'=n$ and $n \pm 1$.

(b) According to the consideration iv), we separate the surface rigidity of the neutron or proton core into two parts:

$$C = C^{(1)} + C^{(2)}, \quad (3 \cdot 13)$$

where $C^{(2)}$ depends strongly on the level spacing (the strongly shell dependent part), and $C^{(1)}$ depends rather weakly on it:

$$C^{(1)} = (5/2\pi) \cdot \sum_{nljm}^{closed\ shell} E_{nljm}^{(2)'} , \quad C^{(2)} = (5/2\pi) \sum_{nljm}^{closed\ shell} E_{nljm}^{(2)''} \quad (3 \cdot 14)$$

Here $E_{nljm}^{(2)'}$ is the term which does not depend on the near-lying levels, as given by

$$E_{nljm}^{(2)'} = (\hbar^2/2MR_0^2) \omega_{nl}^2 [1 + 2(ljm|P_2(\theta)|ljm)] \\ + \sum_{n'=n-1, n+1}^{n+1} \sum_{l', j'}' |(nljm|H^{(1)}|n'l'j'm)|^2 / (W_{nlj}^0 - W_{n'l'j'}^0), \quad (3 \cdot 15)$$

where the "prime" of \sum' means the exclusions of coupled terms with the near-lying levels. $E_{nljm}^{(2)'}$ is the part which depends strongly on the near-lying levels, as given by

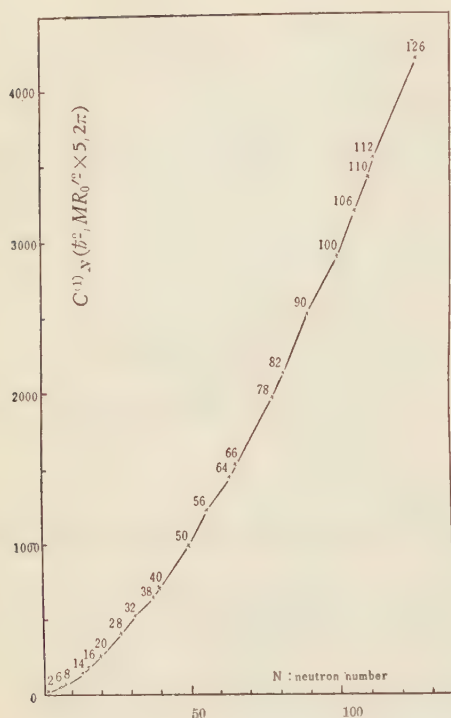
$$E_{nljm}^{(2)''} = \sum_{n'l'j'}'' |(nljm|H^{(1)}|n'l'j'm)|^2 / (\Delta_{nlj-n'l'j'}) \cdot (\hbar^2/2MR_0^2) \\ \Delta_{nlj-n'l'j'} = (W_{nlj}^0 - W_{n'l'j'}^0) / (\hbar^2/2MR_0^2), \quad (3 \cdot 16)$$

where \sum'' means the sum of only the coupled terms with the near-lying levels.

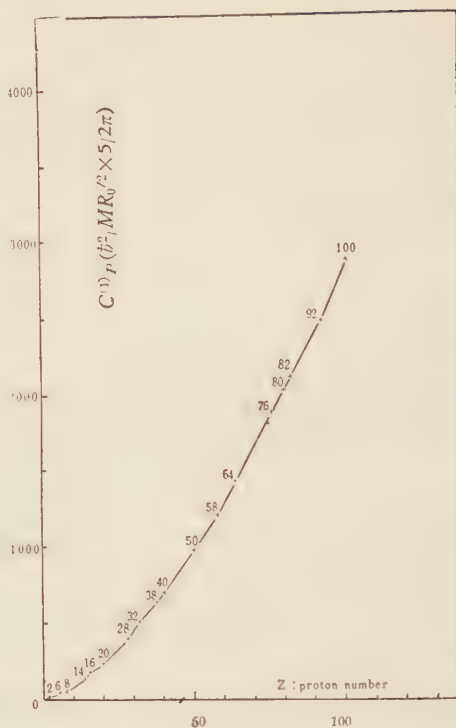
§ 4. Results of calculation and discussions

As discussed above, $C^{(1)}$ does not depend on the spacing between the near-lying levels. In its numerical calculation, therefore, we used for W_{nlj}^1 the expression (3·7) and for g and k in it the values adopted in Fig. I. The result thus obtained is given in Fig. IIa and IIb.

The level ordering of 3s 1/2, 1h 11/2, and 2d 3/2 states and of 2f 5/2, 3p 3/2, 3p 1/2 and 1i 13/2 states are not definite because of the pairing effect, so that, in these regions, the neutron or proton numbers which constitute the closed subshells are uncertain. However, the values of $C^{(1)}$ do not so much depend on the level order of these states. Therefore, we adopted here, according to Mayer and Jensen⁷⁾, the level orderings; 3s 1/2,

Fig. IIa. $C^{(1)}_N$ for neutron cores.

In this figure the Coulomb correction is not included.

Fig. IIb. $C^{(1)}_P$ for proton cores.

In this figure the Coulomb correction is not included.

1h 11/2, 2d 3/2 and 2f 5/2, 3p 3/2, 3p 1/2, 1i 13/2 for neutrons, and 1h 11/2, 2d 3/2, 3s 1/2 for protons.

As is shown in Fig II, the values of $C^{(1)}$ of the closed shells lie on a fairly smooth curve.

Contrary to $C^{(1)}$, $C^{(2)}$ is just the term which results from the coupling between the finally filled levels of the closed shell under consideration and their near-lying levels, so that it depends very strongly on the spacing between these levels.* The fairly smooth curve plotted by the values of $C^{(1)}$ for the closed shells is considerably transfigured by adding $C^{(2)}$ to it.

Now we can give the qualitative discussions on the surface rigidity of the closed shells. As to the more detailed discussions, however, some considerations are still more necessary, because of the following two reasons:

a) Up to here, our calculations have been performed under the condition of the infinite well potential.

*) $C^{(2)}$ depends also very sensitively on the level ordering. Therefore, in the regions of 3s 1/2, 1h 11/2, 2d 3/2 and of 2f 5/2, 3p 3/2, 3p 1/2, 1i 13/2, where the order of these levels is not definite because of the pairing effect, we must check up general tendency of the variation of $C^{(2)}$ by making up the various possible choices of the order of these levels.

b) The values of $C^{(2)}$ depend considerably on the values of level spacing between the finally filled levels of the closed shell under consideration and their near-lying levels.

For these two problems, we took the following steps:

a') For the "spherical" square well potential, the effects of the "finite" wall height can be given approximately by:

$$E_{V_0, R_0} \simeq E_{\infty, R_0} (1 + \lambda_0 / R_0)^{-2} \quad (4.1)$$

$$\lambda_0 = [b^2 / (2MV_0)]^{1/2},$$

as obtained by Feenberg-Hammack. Namely, in this case, we can take into account the effect of the "finite" wall height by replacing only the nuclear radius R_0 by

$$R_0' = R_0 + \lambda_0. \quad (4.2)$$

In the case of the spheroidal potential, this procedure might also be possible^{*)}. As a matter of convenience, we adopt here this procedure as the first approximation, since our calculations have been performed in particular under the condition that the deformation of the potential is small.

Table 1.** The values of level spacings in $\hbar^2/2MR_0'^2$ unit used for our calculation of $C^{(2)}$.

| | 2s1/2 | 1d3/2 | 2p3/2 _e | 1f5/2 | 2p1/2 | 2d5/2 | 1g7/2 | 3s1/2 | 2d3/2 | 2f7/2 | 1h9/2 | 2f5/2 | 3p3/2 | 3p1/2 | 2g9/2 |
|--------|-------|-------|--------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1d5/2 | 5.21 | | | | | | | | | | | | | | |
| 2s1/2 | | 5.03 | | | | | | | | | | | | | |
| 1f7/2 | | | 6.59 | | | | | | | | | | | | |
| 2p3/2 | | | | 5.26 | | | | | | | | | | | |
| 1f5/2 | | | | | 5.03 | | | | | | | | | | |
| 1g7/2 | | | | | | 10.4 | | | | | | | | | |
| 2d5/2 | | | | | | | 2.66 | 7.58 | | | | | | | |
| 1g7/2 | | | | | | | | | 6.73 | | | | | | |
| 3s1/2 | | | | | | | | | 1.99 | | | | | | |
| 1h11/2 | | | | | | | | | | 15.8 | | | | | |
| 2f7/2 | | | | | | | | | | | 1.63 | | 7.26 | | |
| 1h9/2 | | | | | | | | | | | | 6.99 | | | |
| 2f5/2 | | | | | | | | | | | | | 2.71 | 4.51 | |
| 1i13/2 | | | | | | | | | | | | | | | 13.0 |

*) Strictly speaking, it is necessary to investigate to what extent this approximation is possible in the deformed potential.

**) In this table, the value of $\Delta_{2s1/2-1d5/2}$ is larger than that used ordinarily. In our calculation, however, we used the values which are deduced from the separation energy, because in our case Δ means the energy difference between the ground state of the corresponding closed shell and the state of the single particle excitation.

b') The values of level spacing between the finally filled levels of the closed shell under consideration and their near-lying levels, $J_{n l j - n' l' j'}$, are very important in evaluating $C^{(2)}$. So we should take those as close as possible to real values rather than those obtained from the level scheme given in Fig. I. Hence, we adopt those (in unit of $\hbar^2/2MR_0'^2$) which are deduced from the separation energy in the vicinity of nuclei having the corresponding neutron or proton core. And in the case where this deduction is not possible, we adopt those (in unit of $\hbar^2/2MR_0'^2$) used by Horie-Arima.¹⁸ The values of level spacing thus obtained are listed in Table 1. Here we used as the nuclear radius, $R_0 = 1.4 \times 10^{-13} A^{1/3}$ cm, and as the depth of potential, $V_0 = 36$ Mev. And for the values of mass number A we adopted those of the nuclei having the corresponding (neutron or proton) cores.

The surface rigidity of the closed shell thus calculated, $C = C^{(1)} + C^{(2)}$, is shown in Fig. IIIa and IIIb.

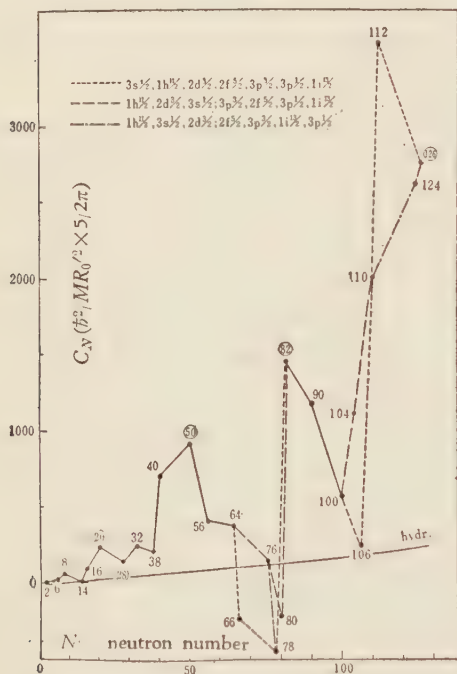


Fig. IIIa The surface rigidity of neutron cores.

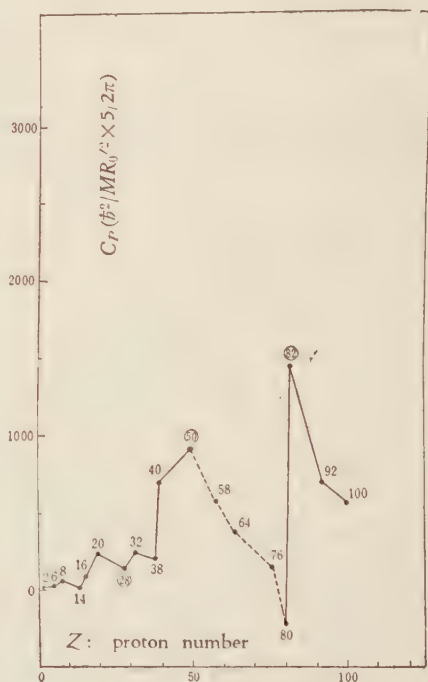


Fig. IIIb The surface rigidity of proton cores.

In this figure the Coulomb correction is not included. Here $R_0' = R_0 + \lambda_0$, $\lambda_0 = (\hbar^2/MV_0)^{1/2}$, where R_0 is the nuclear radius and V_0 is the depth of the potential. The dotted lines are used for the region in which the assignment of subshells is uncertain because of the pairing effect. The lines joining the points for the closed shells have no meaning. In these regions, the rigidity may be reduced very much because of the crossing of the particle levels due to the large deformation of the core which is caused by the coupling of the core and extra-particles. The line indicated by "hydr." is calculated by using the formula

$$C_N^{\text{hydr.}} = (4R_0^2 S) \times N/A,$$

where S is the hydrodynamical surface tension,

$$4\pi R_0^2 S = 15.4 \cdot A^{2/3} \text{ Mev.}$$

Here the dotted lines are used for the region in which the assignment of subshells is uncertain because of the pairing effect. In this region, we calculated C by making up the three possible choices of level ordering.

From Fig. III one can observe a rapid variation of C from nucleus to nucleus, which has never been obtained with the hydrodynamical model. Its qualitative feature are summarized as follows:

- i) The values of the surface rigidity of the closed shells corresponding to the magic numbers are very large except for 28 shell.
- ii) The values of the surface rigidity of the subshells between the neutron or proton numbers 50 and 82 become very small, irrespective of the various possible choices of the level ordering.
- iii) The values of the surface rigidity of the subshells between the neutron or proton numbers 82 and 126 are relatively small except for a particular choice of the level ordering.

The conclusion i), which has been asserted intuitively in connection with the shell model, means that the weak coupling scheme of the collective model applies best to the nuclei in the vicinity of the major closed shells. The conclusions ii) and iii) mean that the strong coupling scheme of the collective model is successful in these regions.

In addition to this general tendency, there is a relatively irregular variation of the rigidity of each closed shell. In comparison with the electric quadrupole moments of the "core \pm one extra-particle type" nuclei, we shall discuss this variation in the next section in detail.

§ 5. Nuclear quadrupole moments and variations of surface rigidity of cores

The magnitudes of the electric quadrupole moments directly reveal their collective origin and the empirical values of quadrupole moments give the direct information about the rigidity of nuclear cores. Therefore, the characteristic variation of the rigidity of cores, which is shown in Fig. III, should be reflected in the quadrupole moments of the "core \pm one extra-particle type" nuclei.

According to the collective model, the quadrupole moments of the "core \pm one extra-particle type" nuclei are given by

$$Q = Q_{s,p} + Q_s, \quad (5.1)$$

of which the first part is due to the extra particle outside the core. The second part is due to the deformed core and is given by:

$$Q_s = P_Q(x) Q_0$$

$$Q_0 = -\frac{3}{4\pi} \frac{2I-1}{2(I+1)} \frac{K}{C} Z R_0^2$$

$$P_Q(x) = 1 - 3 \frac{2I+1}{(I+1)(2I+3)} \frac{x^2}{\sqrt{x^4+4/9}} \quad (\text{for } I=j \gg 1),$$

where Q_0 is the intrinsic quadrupole moment of the core relative to the axes fixed in the core. Here K is the expectation value of the interaction term $H_{int}(x_p, \alpha)$ of (2.13) with respect to the radial part of particle coordinates. Correctly, its value depends on the shell structure, but here we adopt 40 Mev according to B-M's^{*)}. $P_Q(x)$ is referred to as the projection factor. x is a dimensionless parameter^{**)'} which is related to the strength of the particle-surface coupling and is given explicitly by

$$x = \sqrt{\frac{5}{6\pi}} \frac{1}{\sqrt{j}} \frac{K}{\sqrt{\hbar\omega C}}$$

where

$$\omega = \sqrt{C/B}$$

$$B = \frac{1}{2} \frac{3}{4\pi} AMR_0^2.$$

By using the values of the surface rigidity C given by Fig. III, we can calculate the quadrupole moments of the "core \pm one extra-particle type" nuclei from (5.1) and (5.2). Table II shows the comparison of the calculated values of such quadrupole moments with observed ones and hydrodynamical ones. Here the Coulomb correction for C is assumed to be

$$C_{coul} = - \frac{3}{2\pi} \frac{1}{5} \frac{Z^2 e^2}{R_0}.$$

It is very interesting to see whether or not the variation of the rigidity of neutron or proton cores in Fig. III can explain the qualitative tendency of quadrupole moments of the "core \pm one extra-particle type" nuclei.

a) Rigidity of neutron or proton cores between 8 and 50

In this region, the major closed shells are given at the proton or neutron numbers 8, 20, 28 and 50. As is shown in Fig. III, the values of the rigidity of these cores are remarkably large except for 28-shell. Therefore, the weak surface-particle coupling scheme may be valid for ${}_{8}\text{O}_9^{17}$, ${}_{10}\text{K}_{20}^{39}$ and ${}_{40}\text{Zr}_{51}^{91}$, and the values of the quadrupole moments of these nuclei may be much smaller than those of the hydrodynamical estimation. There are no experimental values of the quadrupole moments of ${}_{19}\text{K}_{20}^{39}$ and ${}_{40}\text{Zr}_{51}^{91}$, but the facts that the observed value of that of ${}_{8}\text{O}_9^{17}$, $Q_{ob.}(\text{O}^{17})$, is -0.005 and the hydro-

*) If an extra-particle is exchanged with a hole having the same single particle state, the sign of K is reversed.

**) For $x \sqrt{j} \ll 1$, the weak coupling scheme is valid, but for $x \gg 1$, the strong coupling scheme is valid.

*** Recently, the quadrupole moment of ${}_{19}\text{K}_{20}^{39}$ was measured by Ritter and Series (Proc. Phys. Soc. 68A, 450). Its value is $0.14 \times 10^{-24} \text{cm}^2 \pm 25\%$ which is much smaller than $Q_{hydr.}$ and is very close to $Q_{coul.}$ The authors are indebted to Profs. Townes and Murakawa for bringing this work to their attention

Table II.

The values of surface rigidity C and of quadrupole moments for the "core \pm one extra-particle type" nuclei.

| Nucleus | Configuration | | I | C (Mev) | | Q ($\times 10^{-24}$ cm ²) | | | |
|------------------------------|-------------------|-------------------|-----|---------|-------|---|--------|-------|---------|
| | proton | neutron | | Cal. | Hydr. | Obs. | Cal. | Hydr. | S.P. |
| $^5\text{B}_6^{11}$ | $(1p_{3/2})^{-1}$ | — | 3/2 | 117 | 24.1 | 0.036 | 0.044 | 0.093 | 0.025 |
| $^{13}\text{Al}_{14}^{27}$ | $(1d_{5/2})^{-1}$ | — | 5/2 | 26 | 38.8 | 0.16 | 0.37 | 0.24 | 0.062 |
| $^{17}\text{Cl}_{20}^{37}$ | $1d_{3/2}$ | — | 3/2 | 348 | 45.6 | -0.062 | -0.087 | -0.25 | -0.051 |
| $^{19}\text{K}_{20}^{39}$ | $(1d_{3/2})^{-1}$ | — | 3/2 | 505 | 45.6 | 0.14 | 0.089 | 0.33 | 0.055 |
| $^{27}\text{Co}_{32}^{59}$ | $(1f_{7/2})^{-1}$ | — | 7/2 | 299 | 54.8 | 0.5 | 0.30 | 0.90 | 0.12 |
| $^{31}\text{Ga}_{36}^{69}$ | $(2p_{3/2})^{-1}$ | — | 3/2 | 323 | 58.6 | 0.23 | 0.20 | 0.52 | 0.08 |
| $^{31}\text{Ga}_{40}^{71}$ | $(2p_{3/2})^{-1}$ | — | 3/2 | 691 | 60.2 | 0.14 | 0.14 | 0.52 | 0.08 |
| $^{37}\text{Rb}_{50}^{87}$ | $(2p_{3/2})^{-1}$ | — | 3/2 | 736 | 64.5 | 0.14 | 0.17 | 0.68 | 0.09 |
| $^{59}\text{Pr}_{82}^{141}$ | $2d_{5/2}$ | — | 5/2 | 973 | 67.0 | -0.05 | -0.35 | -1.7 | -0.18 |
| $^{83}\text{Bi}_{126}^{209}$ | $1h_{9/2}$ | — | 9/2 | 1580 | 58.0 | -0.4 | -0.55 | -5.2 | -0.30 |
| $^8\text{O}_9^{17}$ | — | $1d_{5/2}$ | 5/2 | 174 | 28.6 | -0.005 | -0.031 | -0.14 | -0.0013 |
| $^{16}\text{S}_{17}^{33}$ | — | $1d_{3/2}$ | 3/2 | 188 | 41.2 | -0.055 | -0.059 | -0.20 | 0 |
| $^{16}\text{S}_{19}^{35}$ | — | $(1d_{3/2})^{-1}$ | 3/2 | 350 | 45.6 | 0.038 | 0.036 | 0.2 | 0 |
| $^{32}\text{Ge}_{41}^{73}$ | — | $1g_{9/2}$ | 9/2 | 677 | 60.2 | -0.2 | -0.11 | -1.1 | 0 |
| $^{38}\text{Sr}_{49}^{87}$ | — | $(1g_{9/2})^{-1}$ | 9/2 | 733 | 64.5 | — | 0.14 | 1.4 | 0 |
| $^{40}\text{Zr}_{51}^{91}$ | — | $2d_{5/2}$ | 5/2 | 1030 | 62.2 | — | -0.083 | -0.90 | 0 |

dynamical estimation $Q_{hydr.}(\text{O}^{17})$ is -0.14 support the above argument. The hydrodynamical estimation, $Q_{hydr.}(\text{O}^{17}) = -0.14$, differs from the empirical value by one order of magnitude. Our calculated value $Q_{cal.}(\text{O}^{17}) = -0.031$ is better than $Q_{hydr.}(\text{O}^{17})$, but is too large in its absolute value.

Fig. III indicates that the values of rigidity of 14-shell in both protons and neutrons are smaller than those of closed shells in this region. This suggests that the empirical value of the quadrupole moment of $^{13}\text{Al}_{14}^{27}$ may be fairly different from $Q_{s.p.}(\text{Al}^{27})$. In fact, we can easily understand that the deformation of Al^{27} is fairly large, by the facts that $Q_{ob.}(\text{Al}^{27}) = 0.16$, $Q_{s.p.}(\text{Al}^{27}) = 0.062$ and $Q_{hydr.}(\text{Al}^{27}) = 0.24$. Our calculated value, $Q_{cal.} = 0.37$ is worse than the hydrodynamical value. The examples for O^{17} and Al^{27} seem to indicate that the collective treatment of light nuclei is not legitimate.

Fig. III shows that there is a rapid variation of the surface rigidity between 16-shell and 20-shell. This must be reflected in the quadrupole moment ratio of $^{16}\text{S}_{17}^{33}$ (neutrons: 16-shell+1) and $^{16}\text{S}_{19}^{35}$ (neutrons: 20-shell-1) which have same nuclear spin. It is obvious that we obtain

$$|Q_{hydr.}(\text{S}^{33})|/|Q_{hydr.}(\text{S}^{35})| = 1$$

and

$$Q_{s.p.}(\text{S}^{33}) = Q_{s.p.}(\text{S}^{35}) = 0.$$

Contrary to this, the observed values give a ratio

$$|Q_{ob.}(\text{S}^{33})|/|Q_{ob.}(\text{S}^{35})| = 1.45.$$

We may suppose that this value different from unity is due to the rapid variation of the surface rigidity. The ratio given by our calculation is

$$|Q_{cal.}(S^{33})|/|Q_{cal.}(S^{35})|=1.64.$$

This is the major success of our theory.

The fact $Q_{ob.}(Cl^{37})/Q_{ob.}(Cl^{35})=1.23$ may also imply that the rigidity of the core of Cl^{37} is larger than that of Cl^{35} because of the 20-neutron shell, though the framework of our method forbids to calculate $Q(Cl^{35})$.

Fig. III shows that between 38-shell and 40-shell there is a large variation of the rigidity. This should also be reflected in the quadrupole moment ratio of ${}_{31}Ga_{33}^{69}$ and ${}_{31}Ga_{40}^{71}$. The single particle model and the hydrodynamical one give

$$|Q_{s.p.}(Ga^{69})|/|Q_{s.p.}(Ga^{71})|=|Q_{hydr.}(Ga^{69})|/|Q_{hydr.}(Ga^{71})|=1.$$

According to our calculation, $|Q_{cal.}(Ga^{69})|/|Q_{cal.}(Ga^{71})|=1.4$ because of the variation of the rigidity in these two isotopes. This is compared with the observed one; $|Q_{ob.}(Ga^{69})|/|Q_{ob.}(Ga^{71})|=1.6$. This again shows the success of our theory. Moreover, the absolute values of $Q_{ob.}(Ga^{69})=0.23$ and $Q_{ob.}(Ga^{71})=0.14$ are explainable by our theory, whereas both the single particle and the hydrodynamical models fail to explain them. This fact supports that the rigidity of 40-shell is very large as indicated in Fig. III.

The 28-shell is assigned to a major closed shell. Fig. III indicates, however, that the rigidity of the 28-shell is relatively smaller than that of other major closed shells. Therefore we may suppose that the quadrupole moments of the nuclei having this core may fairly deviate from the single particle moments. Namely, the deformations of such nuclei may be larger than those of the nuclei lying near other major closed shells. In fact, $Q_{ob.}(Co^{59})$ is 0.5 ± 0.2 , which is significantly larger than $Q_{s.p.}(Co^{59}) (=0.12)$, while our calculation gives $Q_{cal.}(Co^{59})=0.3$ in agreement with $Q_{ob.}$ within the experimental error.

b) The rigidity of neutron or proton cores between 50 and 82 and between 82 and 126

In these regions, the numbers of the neutrons or protons which compose the closed subshells are not clear because of the pairing effect. In Fig. III, therefore, we took the several possible cases of the level ordering.

The following tendency is maintained in these regions, irrespective of the choice of level ordering. The rigidity of the 50, 82 and 126 (neutron or proton) shells is extremely large and the rigidity of the subshells between the 50 and 82 major closed shells is very small. And the values of the surface rigidity of the subshells between the 82 and 126 major closed shells are relatively small as compared with that of the major closed shells.

This explains why the strong coupling scheme of the collective model is successful for nuclei of mass numbers between 155 and 185. On this point it should be noted that the rigidity of nuclei with "core + several extra-particles" may be considerably reduced compared with the rigidity of the "spherically symmetric core" because of the crossing of the particle levels due to the deformation of the core which is caused by the extra-particles.

The fact that the rigidity of the 82- and 126-shell is very large can be inferred, as is well known, from $Q_{ob.}(\text{Bi}^{209})$ which is very close to $Q_{s.p.}(\text{Bi}^{209})$.

Fig. IIIb indicates that there is an abrupt change between the values of the rigidity of 80- and 82-proton shell. The fact that the quadrupole moment of ${}_{79}\text{Au}^{197}$ ($=0.56$), is much larger than $Q_{s.p.}(\text{Au}^{197})$ might reflect this situation.

§ 6. Discussions

We have tried to make clear the relation between the rigidity of the core and the proper shell structure of it, and have calculated the rigidity concretely by using a simplified picture, and further have discussed the qualitative tendency in the variation of the rigidity of cores, in comparison with the quadrupole moments of the "core \pm one extra-particle type" nuclei. Then, we have found the results obtained can qualitatively explain the characteristic variation of the rigidity of cores, which must be accepted as an empirical fact. A similar feature is also obtained by Horie and Arima¹⁵⁾ in their method of configuration mixing. Its relation to ours will be discussed in a forthcoming paper.

Of course, the numerical values of our calculation are not quantitative enough, because of the use of the simplified picture where the effects of the direct inter-particle interactions are neglected, and of the adoption of various approximations in the course of calculations. As the general tendency, however, the calculated values of surface rigidity are much larger than those obtained from the hydrodynamical surface tension. On this point it should be emphasized that our calculation is performed only for the closed shells which are of spherical symmetry.

When the core has a large deformation owing to extra-particles, there may occur the crossing of particle levels, which causes the rearrangement of the level filling of particles.²⁰⁾ Therefore, the surface rigidity of such a deformed core may be very much different from that of the spherically symmetric core. Preliminary discussion on this problem was given by Gallone and Salvetti.¹⁹⁾ They showed, by using an extremely simplified model, that the surface rigidity is much reduced by the large deformation and on the average it approaches to the hydrodynamical one.

This can easily be inferred from the fact that the rigidity of the hydrodynamical model is due to a statistical feature of the level distributions. The increment of the crossing of particle levels as the deformation increases may mean that the level distribution characterized by the shell model becomes more and more the statistical one.

In the regions where the strong particle-surface coupling scheme is successful, therefore, the hydrodynamical rigidity may be better in the case of several extra-particles, because in these regions the deformations of cores are considerably large.

More detailed discussions on this point is being prepared together with the analyses of the quadrupole moments of the "core + several extra-particles type" nuclei.

To conclude, we should like express our thanks to Professor S. Hayakawa for his kind discussions and continuous encouragement in the course of our work. This work was performed under the financial aid of "Iwanami Fujukai", and their good will is highly appreciated.

References

- 1) A. Bohr and B. R. Mottelson, Dan. Mat.-fys. Medd. **27**, no. 16 (1953).
- 2) D. L. Hill and J. A. Wheeler, Phys. Rev. **89** (1953), 1102.
- 3) J. Rainwater, Phys. Rev. **79** (1950), 432.
- 4) E. Feenberg and K. C. Hammack, Phys. Rev. **81** (1953), 1102.
- 5) K. W. Ford, Phys. Rev. **90** (1953), 29.
- 6) R. van Wageningen and J. De Boer, Physica **18** (1952), 369.
- 7) M. G. Mayer and J. H. Jensen, "*Elementary theory of nuclear shell structure*" (1955).
- 8) D. Pfrisch, Z. Physik **132** (1952), 409.
- 9) S. Gallone and C. Salvetti, Nuovo Cimento **10** (1953), 145.
- 10) S. A. Moszkowski, Phys. Rev. **99** (1955), 803.
- 11) T. Marumori, J. Yukawa and R. Tanaka, Prog. Theor. Phys. **13** (1955), 442.
T. Marumori and E. Yamada, Prog. Theor. Phys. **13** (1955), 557.
- 12) S. Tomonaga, Prog. Theor. Phys. **13** (1955), 467.
- 13) T. Miyazima and T. Tamura, Prog. Theor. Phys. **15** (1956), 255.
- 14) T. Marumori, Prog. Theor. Phys. **14** (1955), 608.
- 15) H. Horie and A. Arima, Phys. Rev. **99** (1955), 778.
- 16) R. J. Eden and N. C. Francis, Phys. Rev. **97** (1955), 1366.
- 17) H. A. Tolhoek, Physica **21** (1955), 144.
- 18) F. Coester, Phys. Rev. **99** (1955), 170.
- 19) P. F. A. Klinkenberg, Rev. Mod. Phys. **24** (1952), 63.
- 20) B. R. Mottelson and S. G. Nilsson, Phys. Rev. **99** (1955) 1615.

Note added in proof Recently, J. M. Araújo has calculated the nuclear deformability by using the generalized independent particle model (time dependent well potential) [Nuclear Physics, Vol. 1, No. 4 (1956)]. The relation between Araújo's description and ours will be discussed in later occasion.

On the Theory of Condensation^{*,†}

Kazuyosi IKEDA

*Physics Department, Faculty of Science, Kyusyu University,
Fukuoka, Kyusyu*

(Received October 20, 1954)**

The Mayer-Born-Kahn-Uhlenbeck theory of imperfect gases and condensation phenomena has treated the *ideal* systems, for which the volume dependence of the cluster integrals is completely neglected, but the present paper treats the *real* systems, for which the volume dependence of the cluster integrals is rigorously taken into consideration.

We prove a sequence of lemmas and theorems, which may be useful for the treatment of the *real* systems, and which may therefore contribute to the theory of imperfect gases and condensation phenomena. (Part I.)

Then, from the analytical viewpoint, we discuss the condensation phenomena of the *real* systems, and discuss whether Mayer's theory about the connection between the phase transitions and the singularities is correct or not. (Parts II and III.)

Introduction

§ 0.1. Consider a system composed of N identical molecules (particles) of mass m in a volume V and at a temperature T . Then, if we follow classical statistical mechanics, the partition function Q_N of the system is given by

$$Q_N = (2\pi mkT/h^3)^{3N/2} \mathcal{Q}_N, \quad (1)$$

where k and h are Boltzmann's and Planck's constants respectively and \mathcal{Q}_N is the configuration integral given by

$$\mathcal{Q}_N = (1/N!) \iiint_V \dots \exp(-U_N/kT) d\tau_1 d\tau_2 \dots d\tau_N, \quad (2)$$

U_N being the total potential energy of the system. The Helmholtz free energy A of the system is given by

$$A = -kT \ln Q_N = NkT \ln (h/\sqrt{2\pi mkT})^3 - kT \ln \mathcal{Q}_N, \quad (3)$$

from which all the other thermodynamic properties of the system can be derived; in particular, the equation of state can be obtained:

* Translated from the papers^{1), 15)} which appeared in Japanese in 1952 and 1953.

† This paper was presented at the International Conference of Theoretical Physics, Kyôto, September 22, 1953.—Cf. the proceedings of the conference.²⁾

** The original manuscript of this paper was first received about two years ago, but unfortunately was misplaced in our office. At our request, the author kindly undertook to resubmit his manuscript. We wish to express our deep regret for the long delay in the publication of this paper.

$$p = -(\partial A / \partial V) = kT (\partial / \partial V) \ln \Omega_N, \quad (4)$$

p being the pressure. Now we have for Ω_N the expression

$$\Omega_N = \sum_{m_l \geq 0} \left(\sum_{l=1}^N l m_l = N \right) \prod_{l=1}^N \frac{(V b_l)^{m_l}}{m_l!}, \quad (5)$$

which is called the Ursell-Mayer development.^{(3),(4)} Here the b_l 's are the so-called "cluster integrals", each of which depends on the volume V and the temperature T as well as on the intermolecular forces.

Using the expression (5), Mayer and others^{(4),(5),(6)} developed a very elegant theory of imperfect gases and condensation phenomena. But in the Mayer theory the volume dependence of the cluster integrals was completely neglected. That is to say, by Born and Fuchs⁽⁵⁾ and Kahn and Uhlenbeck⁽⁶⁾ (who improved Mayer's original theory mathematically) as well as by Mayer⁽⁴⁾, the "(0)-system", a sort of an *ideal* system, was treated, where the configuration integral of the "(0)-system" is defined to be

$$\Omega_N^{(0)}(V, T) = \sum_{m_l \geq 0} \left(\sum_{l=1}^N l m_l = N \right) \prod_{l=1}^N \frac{\{V b_l^{(0)}(T)\}^{m_l}}{m_l!}, \quad (6)$$

the $b_l^{(0)}(T)$'s being the volume-independent cluster integrals (which we shall call the non-boundary cluster integrals), $b_l^{(0)}(T) = \lim_{V \rightarrow \infty} b_l(V, T)$. Then the limit $N \rightarrow \infty$, $V = Nv \rightarrow \infty$ with v (volume per molecule) fixed, was taken, and the limiting function

$$\omega^{(0)}(v, T) = \lim_{N \rightarrow \infty} (1/N) \ln \Omega_N^{(0)}(Nv, T) \quad (7)$$

was used to derive the thermodynamic properties (per molecule), and the singularity of the function $\omega^{(0)}$ of v , farthest from the origin of v , was considered to give the point of condensation.

However, in reality, we should treat not the "(0)-system" but the "real system", i.e., the system for which the volume dependence of the cluster integrals is rigorously taken into consideration; the configuration integral of the "real system" is

$$\Omega_N^\dagger(V, T) = \sum_{m_l \geq 0} \left(\sum_{l=1}^N l m_l = N \right) \prod_{l=1}^N \frac{\{V b_l^\dagger(V, T)\}^{m_l}}{m_l!}. \quad (8)$$

Here we have used the superscript \dagger to emphasize that this system is "real". The $b_l^\dagger(V, T)$'s are the *real* rigorously volume-dependent cluster integrals. Taking the limit $N \rightarrow \infty$, $V = Nv \rightarrow \infty$ with v fixed, we should use the limiting function

$$\omega^\dagger(v, T) = \lim_{N \rightarrow \infty} (1/N) \ln \Omega_N^\dagger(Nv, T) \quad (9)$$

to derive the thermodynamic properties and to determine the point of condensation.

§ 0.2. The purposes of the present paper are as follow:

(i) We treat the "real system"; that is, we discuss the thermodynamic properties and the condensation phenomenon of the "real system."

(ii) With a view to doing this, we seek theorems which give conditions for the truth of the equality

$$\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^i(Nv, T) = \lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^{(0)}(Nv, T) \quad (10)$$

i.e. $\omega^+(v, T) = \omega^{(0)}(v, T).$

For, as long as this equality is true, the thermodynamic properties of the "real system" are equal to those of the "(0)-system", that is, the "real system" behaves thermodynamically in exactly the same way as the "(0)-system". (Part I.)

(iii) Using some of the theorems obtained, we argue about the condensation phenomenon of the "real system" and about the analytical properties of the singular point with which the condensation is connected; we thus discuss whether Mayer's theory about the connection between the phase transitions and the singularities is correct. (Parts II and III.)

§ 0.3. The arguments in the present paper have the following features:

(i) We argue from the analytical viewpoint. A macroscopic system (for which $N \simeq 10^{23}$ in the physical sense) is regarded as an infinite system (for which $N \rightarrow \infty$ in the mathematical sense), and so the concept of "limit" in analysis may be used here; thus the phase transitions are connected with the singularities (in the mathematical sense) of the functions representing the thermodynamic properties.

(ii) We argue by a rigorous mathematical method. All of the equalities and inequalities appearing in Part I of the present paper are to be read in the pure mathematical sense*. We believe that such a method will be essential for the analytical discussion of phase transitions.

Part I

Theory of Imperfect Gases with Volume-Dependent Cluster Integrals

§ 1.0. In this article we shall prove a sequence of lemmas and theorems, which may give a theory of imperfect gases with volume-dependent cluster integrals, and which may also be utilized for the theory of condensation of such gases.

In Definition 1, we define the *real cluster integrals* $b_i^+(V, T)$, of which the volume dependence is rigorously taken into consideration, and the *non-boundary cluster integrals* $b_i^{(0)}(T)$, of which the volume dependence is completely neglected; in addition we define three sorts of integrals, which describes the behaviour of cluster integrals near the boundary of the vessel (assumed to be cubic). Then, in Lemmas 1, 2 and 3 (with Definition 2), we discuss to what extent the $b_i^+(V, T)$'s deviate from the $b_i^{(0)}(T)$'s by the boundary effect or volume effect. (§ 1.1.)

* In our arguments we of course distinguish, as we should in mathematics, between the *wahrhaft* infinity and the *uneigentlich* infinity; the former appears in case of "an infinite number of..." or "an infinite set (or sequence)" and the latter appears in case of "an infinite limit-value" or "infinity" (which is an ideal number). (cf. the theory of sets.)

In Lemma 4, we prove the existence of a function $l^*(N)$ which satisfies certain conditions (1'–5'). In general there exist many such $l^*(N)$'s for any given T and v ; but in Definition 3 we take any one of these. We use this $l^*(N)$ to discriminate between *large* and *small* clusters [large $l > l^*(N)$, small $l \leq l^*(N)$]. In a word, $l^*(N)$ is the border-line between "large" and "small", and it plays an important rôle in the present theory. In fact, we can choose, and have chosen, $l^*(N)$ suitable (i.e., increasing neither too rapidly nor too slowly with N) so that Lemmas 5 and 6 hold. In Definition 4, for any given T and v we define the infinite sequences of "real systems" and "(0)-systems" ($N=1, 2, \dots$ to ∞) and express the partition function of each system (for almost every N) as the sum of two parts, viz. the *small cluster part*, each term of which contains no *large* cluster, and the *large cluster part*, each term of which contains at least one *large* cluster. (§ 1.2.)

Then, using Lemmas 3 and 4, we prove Lemma 5, which states equality (in the limit of infinite system) between the small cluster part of the partition function for the "real system" and that for the "(0)-system". (§ 1.3.)

Then, using Lemma 4 and the known properties of the "(0)-system", we prove Lemma 6, which states that for the "(0)-system" and for $v > v_c(T)$ (gaseous phase) the small cluster part alone accounts for the whole partition function. (§ 1.4.)

Then, using Lemmas 5 and 6, we prove Theorem α and Corollaries 1,2 which give conditions that the limit (9) exist and the equality (10) be true, and from which one knows conditions for equality between the thermodynamic properties of the "real system" and of the "(0)-system". (§ 1.5.)

Also we obtain the similar theorem for inequality (Theorem β). (§ 1.6.)

§ 1.1. Volume dependence of the cluster integrals

DEFINITION 1. Let the potential energy of a pair of particles be given by a function $u(r)$ of the distance r (≥ 0) between the pair of particles, and let the distance between the i -th and the j -th particle in a set of particles be indicated by the variable r_{ij} . Using the function

$$f_{ij} \equiv f(r_{ij}, T) = \exp \{-u(r_{ij})/kT\} - 1 \quad (1)$$

(where k is Boltzmann's constant and T is the absolute temperature), we define five kinds of integrals as below. Here the symbol $\sum_{i,j,\dots}^{cluster}$ means the sum over all the products consistent with a single cluster consisting of l particles, where the definition of a "cluster" has been given by Mayer and others.¹⁾⁵⁾⁶⁾ These integrals are given as functions of T (> 0), V (> 0) and x_l (> 0), etc., for every positive integer l , that is to say, as infinite sequences of functions ($l=1, 2, 3, \dots$ to ∞).

(i) *The real cluster integrals*; ‡

‡ The value of the cluster integral $b_l(V, T)$ depends not only on the variables l , V and T but also on the shape of the domain over which the integral extends, that is, the shape of the vessel in which the particles composing the cluster are confined. But here we have specified the shape to be cubic. Then we denote this specification by the superscript \dagger . On the contrary, $b_l^{(0)}(T)$ never concerns any finite domain.

$$b_l^\dagger(V, T) = \frac{1}{l!V} \int_V \sum_{\substack{\text{cluster} \\ l \geq i > j \geq 1}} \prod f_{ij} d\{l\}.$$

(Here the symbol $\int_V d\{l\}$ means $\int_V \dots \int_V d\tau_1 d\tau_2 \dots d\tau_l$.) The integration, with respect to each of the l particles, is extended over a cubic domain of volume V .

(ii) *The non-boundary cluster integrals:*

$$b_l^{(0)}(T) = \frac{1}{l!} \int_{\mathbf{D}^{(0)}} \sum_{\substack{\text{cluster} \\ l \geq i > j \geq 1}} \prod f_{ij} d\{l-1\}.$$

(Here the symbol $\int_{\mathbf{D}^{(0)}} d\{l-1\}$ means $\int_{\mathbf{D}^{(0)}} \dots \int_{\mathbf{D}^{(0)}} \int_{\mathbf{D}^{(0)}} d\tau_1 d\tau_2 \dots d\tau_{l-1}$.) The l -th particle being fixed at a point, the integration, with respect to each of the other $l-1$ particles, is extended over the whole space (with no boundary).

(iii) *The one-boundary cluster integrals:*

$$b_l^{(1)}(x_l, T) = \frac{1}{l!} \int_{\mathbf{D}^{(1)}(x_l)} \sum_{\substack{\text{cluster} \\ l \geq i > j \geq 1}} \prod f_{ij} d\{l-1\}.$$

The l -th particle being fixed at a point distant x_l from a plane, the integration, with respect to each of the other $l-1$ particles, is extended over the half space $\mathbf{D}^{(1)}(x_l)$ (containing the position of the l -th particle) bounded by the plane.

(iv) *The two-boundary cluster integrals:*

$$b_l^{(2)}(x_l, y_l, T) = \frac{1}{l!} \int_{\mathbf{D}^{(2)}(x_l, y_l)} \sum_{\substack{\text{cluster} \\ l \geq i > j \geq 1}} \prod f_{ij} d\{l-1\}.$$

The l -th particle being fixed at a point distant x_l and y_l from two planes which are perpendicular to each other, the integration, with respect to each of the other $l-1$ particles, is extended over the one-fourth space $\mathbf{D}^{(2)}(x_l, y_l)$ (containing the position of the l -th particle) bounded by those two planes.

(v) *The three-boundary cluster integrals:*

$$b_l^{(3)}(x_l, y_l, z_l, T) = \frac{1}{l!} \int_{\mathbf{D}^{(3)}(x_l, y_l, z_l)} \sum_{\substack{\text{cluster} \\ l \geq i > j \geq 1}} \prod f_{ij} d\{l-1\}.$$

The l -th particle being fixed at a point x_l, y_l and z_l distant from three planes which are perpendicular to each other, the integration, with respect to each of the other $l-1$ particles, is extended over the one-eighth space $\mathbf{D}^{(3)}(x_l, y_l, z_l)$ (containing the position of the l -th particle) bounded by those three planes.

Especially for $l=1$ we define

$$b_1^\dagger(V, T) \equiv b_1^{(0)}(T) \equiv b_1^{(1)}(x_1, T) \equiv b_1^{(2)}(x_1, y_1, T) \equiv b_1^{(3)}(x_1, y_1, z_1, T) \equiv 1. \quad (2)$$

Postulate 1. (concerning the function $u(r)$ —the pair potential.)

(i) There exists a finite number $r_0(>0)$ such that $u(r)=0$ if $r \geq r_0$.

(ii) $u(r)$ is lower bounded for $r \geq 0$.

(iii) $f(r) = \exp\{-u(r)/kT\} - 1$ ($T>0$) is continuous for $r \geq 0$ except for at most a finite number of points.

Lemma α . For any given value of l , let K_l be a spherical domain of radius lr_0 whose center is the point at which the l -th particle is fixed. Then we have for any domain D

$$\int_D \sum_{l \geq i > j \geq 1}^{\text{cluster}} f_{ij} d\{l-1\} = \int_{D \cap K_l} \sum_{l \geq i > j \geq 1}^{\text{cluster}} f_{ij} d\{l-1\},$$

where $D \cap K_l$ means the intersection of D and K_l .

[For, by (i) of Postulate 1, in order that at least one term of the sum $\sum f_{ij}$ may be non-zero, it is necessary that all of the $l-1$ particles should lie inside K_l .]

Corollary 1. For any regions D_1, D_2, \dots such that $D_1 \cap K_l = D_2 \cap K_l = \dots$, we have

$$\int_{D_1} * = \int_{D_2} * = \dots = \int_{D_1 \cap K_l} * (* \equiv \sum_{l \geq i > j \geq 1}^{\text{cluster}} f_{ij} d\{l-1\}).$$

Corollary 2. For the whole space $D^{(0)}$ and any regions D', D'', \dots containing K_l (that is, $D' \supset K_l, D'' \supset K_l, \dots$), we have

$$\int_{D^{(0)}} * = \int_{D'} * = \int_{D''} * = \dots = \int_{K_l} * (* \equiv \sum_{l \geq i > j \geq 1}^{\text{cluster}} f_{ij} d\{l-1\}).$$

Corollary 3.

- (i) $b_l^{(1)}(x_l, T) = b_l^{(0)}(T)$ if $x_l \geq lr_0$.
- (ii) $b_l^{(2)}(x_l, y_l, T) = b_l^{(1)}(x_l, T)$ if $y_l \geq lr_0$.
- (iii) $b_l^{(2)}(x_l, y_l, T) = b_l^{(0)}(T)$ if $x_l \geq lr_0$ and $y_l \geq lr_0$.
- (iv) $b_l^{(3)}(x_l, y_l, z_l, T) = b_l^{(2)}(x_l, y_l, T)$ if $z_l \geq lr_0$.
- (v) $b_l^{(3)}(x_l, y_l, z_l, T) = b_l^{(1)}(x_l, T)$ if $y_l \geq lr_0$ and $z_l \geq lr_0$.
- (vi) $b_l^{(3)}(x_l, y_l, z_l, T) = b_l^{(0)}(T)$ if $x_l \geq lr_0, y_l \geq lr_0$, and $z_l \geq lr_0$.

[Hence, it is in fact sufficient to define $b_l^{(1)}(x_l, T)$ in the interval $0 \leq x_l \leq lr_0$, $b_l^{(2)}(x_l, y_l, T)$ in the square interval $0 \leq x_l \leq lr_0, 0 \leq y_l \leq lr_0$, and $b_l^{(3)}(x_l, y_l, z_l, T)$ in the cubic interval $0 \leq x_l \leq lr_0, 0 \leq y_l \leq lr_0, 0 \leq z_l \leq lr_0$.]

Lemma β . (i) The integral $b_l^{(0)}(T)$ is definite and finite for any value of $T(>0)$ and for any value of l .

(ii) The functions $b_l^{(1)}(x_l, T)$, $b_l^{(2)}(x_l, y_l, T)$ and $b_l^{(3)}(x_l, y_l, z_l, T)$ are uniquely defined and bounded in the regions $(x_l \geq 0)$, $(x_l \geq 0, y_l \geq 0)$ and $(x_l \geq 0, y_l \geq 0, z_l \geq 0)$ respectively, for any value of $T(>0)$ and for any value of l .

[For, by (ii) of Postulate 1, for any $T(>0)$ there exists a constant A (independent of r) such that $|f(r)| \leq A$ for $r \geq 0$, and so the integrands $\sum f_{ij}$ are bounded for the coordinates of the particles, while by Lemma α the integration regions $D^{(0)}, D^{(1)}, D^{(2)}$ and $D^{(3)}$ may be replaced by the finite regions $K_l, D^{(1)} \cap K_l, D^{(2)} \cap K_l$ and $D^{(3)} \cap K_l$ respectively with the values of the integrals $b_l^{(0)}, b_l^{(1)}, b_l^{(2)}$ and $b_l^{(3)}$ unchanged; hence by (iii) of Postulate 1 the integrals are definite and finite, since the conditions for integrability are satisfied. Now we have

$$|(1/l!) \int_{\mathbf{D} \sim \mathbf{K}_l} \sum H f_{ij} d\{l-1\}| \leq (1/l!) \{(4\pi/3) (lr_0)^3\}^{l-1} \sum H A_{ij} (A_{ij} - A),$$

which proves $b_l^{(1)}$, $b_l^{(2)}$ and $b_l^{(3)}$ to be bounded for $x_l, y_l, z_l \geq 0$ if we replace \mathbf{D} by $\mathbf{D}^{(1)}$, $\mathbf{D}^{(2)}$ and $\mathbf{D}^{(3)}$.]

Postulate 2. (concerning the cluster integrals.)

(i) For any value of $T(>0)$, there exists a constant $G^{(0)}(>0)$, independent of l , such that $|b_l^{(0)}(T)|^{1/l} \leq G^{(0)}$ for all l .

(ii) For any value of $T(>0)$, there exist constants $G^{(1)}(>0)$, $G^{(2)}(>0)$ and $G^{(3)}(>0)$, independent of l , x_l , y_l and z_l , such that

$$(a) \quad |b_l^{(1)}(x_l, T)|^{1/l} \leq G^{(1)} \text{ for all } l \text{ and } x_l,$$

$$(b) \quad |b_l^{(2)}(x_l, y_l, T)|^{1/l} \leq G^{(2)} \text{ for all } l, x_l \text{ and } y_l,$$

$$(c) \quad |b_l^{(3)}(x_l, y_l, z_l, T)|^{1/l} \leq G^{(3)} \text{ for all } l, x_l, y_l \text{ and } z_l.$$

[This represents the fact that the contribution per particle to any cluster integral remains finite even if the number of the particles composing the cluster integral becomes indefinitely large. In particular, (i) is necessary and sufficient in order that the radius of convergence of the power series $\sum_{l=1}^{\infty} b_l^{(0)}(T) z^l$ should not be zero, since by Cauchy-Hadamard's theorem the radius of convergence, $R(T)$, is given by

$$1/R(T) = \lim_{l \rightarrow \infty} \sup |b_l^{(0)}(T)|^{1/l}. \quad (1)$$

According to the theory of condensation, this is a necessary condition that the condensation of a gas should not occur at the density zero, i.e., the theory itself should not be meaningless.]

$$\text{LEMMA 1. If} \quad 2lr_0/V^{1/3} \leq 1, \quad (1)$$

then we have

$$|b_l^{(1)}(V, T) - b_l^{(0)}(T)| \leq \theta(l, T) \cdot 8lr_0/V^{1/3}, \quad (2)$$

where

$$\theta(l, T) = \max \{ \theta^{(1)}(l, T), \theta^{(2)}(l, T), \theta^{(3)}(l, T) \} \quad (3)$$

$$\theta^{(1)}(l, T) = \text{l.u.b.}_{0 \leq x_l \leq lr_0} |b_l^{(1)}(x_l, T) - b_l^{(0)}(T)| \quad (3a)^*$$

$$\theta^{(2)}(l, T) = \text{l.u.b.}_{\substack{0 \leq x_l \leq lr_0 \\ 0 \leq y_l \leq lr_0}} |b_l^{(2)}(x_l, y_l, T) - b_l^{(0)}(T)| \quad (3b)$$

$$\theta^{(3)}(l, T) = \text{l.u.b.}_{\substack{0 \leq x_l \leq lr_0 \\ 0 \leq y_l \leq lr_0 \\ 0 \leq z_l \leq lr_0}} |b_l^{(3)}(x_l, y_l, z_l, T) - b_l^{(0)}(T)|, \quad (3c)$$

* Throughout this paper, the *least upper bound* (l.u.b.), as well as the *greatest limit* (lim sup), and the *greatest lower bound* (g.l.b.), as well as the *least limit* (lim inf), are allowed to take not only real values but also ideal values $+\infty$ and $-\infty$. Hence lim sup, lim inf, l.u.b. and g.l.b. are all existent in any infinite set of real numbers.

and we have

$$0 \leq \theta(l, T) < +\infty \quad (4)$$

for every finite value of l and for every finite value of $T (> 0)$.

<Proof>—By Lemma β , for any given l and T , $|b_l^{(i)}(x_i, T) - b_l^{(i)}(T)|$ etc. are bounded in the intervals specified here. Hence

$$0 \leq \theta^{(i)}(l, T) < +\infty \quad (i=1, 2, 3). \quad (5)$$

Hence (4). Now assume (1), and then, for each of the six planes forming the cubic domain of volume V , set a plane parallel to it at a distance of lr_0 from it inside the cubic domain. Then by the new six planes the cubic domain is divided into twenty-seven (or twenty-six) rectangular domains, viz. six domains $lr_0 \times (V^{1/3} - 2lr_0) \times (V^{1/3} - 2lr_0)$, whose union is denoted by \mathcal{A}_1 , and twelve domains $lr_0 \times lr_0 \times (V^{1/3} - 2lr_0)$, whose union is denoted by \mathcal{A}_2 , and eight (cubic) domains $lr_0 \times lr_0 \times lr_0$, whose union is denoted by \mathcal{A}_3 , and one (cubic) domain \mathcal{A}_0 i.e. $(V^{1/3} - 2lr_0) \times (V^{1/3} - 2lr_0) \times (V^{1/3} - 2lr_0)$ surrounded by all other domains; especially when equality in (1) holds, the last one vanishes. Then the integral $(1/l!) \int_V \sum_{i=1}^l f_i d\{l-1\}$, which extends over the cubic domain of volume V with respect to each of the $l-1$ particles with the l -th particle fixed at a point of the cubic domain, is equal to $b_l^{(0)}(T)$ if the l -th particle is at a point of \mathcal{A}_0 (by Corol. 2 of Lemma α), and is equal to $b_l^{(1)}(x_i, T)$ if the l -th particle is at a point of \mathcal{A}_1 distant x_i from the nearest wall, and is equal to $b_l^{(2)}(x_i, y_i, T)$ if the l -th particle is at a point of \mathcal{A}_2 distant x_i and y_i from the nearest two walls, and is equal to $b_l^{(3)}(x_i, y_i, z_i, T)$ if the l -th particle is at a point of \mathcal{A}_3 distant x_i, y_i and z_i from the nearest three walls, (by Corol. 1 of Lemma α). Therefore we have

$$\begin{aligned} Vb_l^\dagger(V, T) = & Vb_l^{(0)}(T) + \int_{\mathcal{A}_1} \{b_l^{(1)}(x_i, T) - b_l^{(0)}(T)\} d\tau_i \\ & + \int_{\mathcal{A}_2} \{b_l^{(2)}(x_i, y_i, T) - b_l^{(0)}(T)\} d\tau_i + \int_{\mathcal{A}_3} \{b_l^{(3)}(x_i, y_i, z_i, T) - b_l^{(0)}(T)\} d\tau_i. \end{aligned} \quad (6)$$

Hence

$$\begin{aligned} |b_l^\dagger(V, T) - b_l^{(0)}(T)| & \leq (1/V) \{\mathcal{A}_1 \theta^{(1)}(l, T) + \mathcal{A}_2 \theta^{(2)}(l, T) + \mathcal{A}_3 \theta^{(3)}(l, T)\} \\ & \leq (1/V) (\mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3) \theta(l, T), \end{aligned} \quad (7)$$

where \mathcal{A}_1 , \mathcal{A}_2 and \mathcal{A}_3 mean the volumes of \mathcal{A}_1 , \mathcal{A}_2 and \mathcal{A}_3 respectively. And we have

$$\begin{aligned} 0 \leq \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 & = 6lr_0 V^{2/3} - 12l^2 r_0^2 V^{1/3} + 8l^3 r_0^3 \\ & \leq 6lr_0 V^{2/3} + 8l^3 r_0^3 \leq 8lr_0 V^{2/3}, \end{aligned} \quad (8)$$

by (1). By (7) and (8) we prove (2).

Corollary. $\lim_{V \rightarrow \infty} b_l^\dagger(V, T) = b_l^{(0)}(T)$ for any fixed l and $T (> 0)$.

LEMMA 2. Let $b_l^{(i)}(T) \neq 0$. (1)

Then, if $2lr_0/V^{1/3} \leq 1$ (2)

and $\theta(l, T) \cdot 8lr_0/V^{1/3} \leq \alpha,$ (3)

where $\theta(l, T) = \theta(l, T)/|b_l^{(0)}(T)|$ (4)

and α is a constant such that $0 < \alpha < 1$, then we have

(i) $0 < 1 - \alpha \leq b_l^+(V, T)/b_l^{(0)}(T) \leq 1 + \alpha < 2,$

(ii) $\left| \frac{1}{l} \ln \frac{b_l^+(V, T)}{b_l^{(0)}(T)} \right| \leq \theta(l, T) \cdot \frac{8r_0}{V^{1/3}(1-\alpha)}.$

⟨Proof⟩—Put $\eta(l, V, T) = b_l^+(V, T)/b_l^{(0)}(T) - 1.$ (5)

Then by Lemma 1 we have from (2)

$$|\eta(l, V, T)| \leq \theta(l, T) \cdot 8lr_0/V^{1/3}. \quad (6)$$

Hence by (3)

$$|\eta(l, V, T)| \leq \alpha < 1, \quad (7)$$

which proves (i); and, since $\ln(1+\eta) = \sum_{\nu=1}^{\infty} (-1)^{\nu-1} \eta^{\nu}/\nu$ ($-1 < \eta \leq 1$), we have by (7) and (6)

$$\begin{aligned} |\ln\{1+\eta(l, V, T)\}| &\leq \sum_{\nu=1}^{\infty} |\eta(l, V, T)|^{\nu} \\ &\leq \{\theta(l, T) \cdot 8lr_0/V^{1/3}\} / \{1 - \theta(l, T) \cdot 8lr_0/V^{1/3}\}, \end{aligned} \quad (8)$$

which proves (ii), by (3)

DEFINITION 2. We define a function $\bar{\theta}(l, T)$ of $l (= 1, 2, 3, \dots)$ as

$$\bar{\theta}(l, T) = \max_{l' \leq l} \theta(l', T).$$

Lemma 7. (i) $\bar{\theta}(l, T)$ is a monotone increasing function of l ; hence $\lim_{l \rightarrow \infty} \bar{\theta}(l, T)$ exists and is finite or $+\infty$.

(ii) (a) $\bar{\theta}(l, T) \geq \theta(l', T)$ for $l > l' \geq 1,$

(b) $\bar{\theta}(l, T) \geq \theta(l, T)$ for $l \geq 1.$

(iii) (a) $\lim_{l \rightarrow \infty} \bar{\theta}(l, T) = \text{l.u.b.}_{l \geq 1} \theta(l, T),$

(b) $\text{l.u.b.}_{l \geq l_0} \{\theta(l, T)/\bar{\theta}(l, T)\} = 1$

(where l_0 is the least l for which $\bar{\theta}(l, T) > 0$). That is to say:—

1° If $\max_{l \geq 1} \theta(l, T)$ exists, then

(a') $\lim_{l \rightarrow \infty} \bar{\theta}(l, T) = \max_{l \geq 1} \theta(l, T),$

(b') $\max_{l \geq l_0} \{\theta(l, T)/\bar{\theta}(l, T)\} = 1.$

2° If $\max_{l \geq 1} \theta(l, T)$ does not exist, then

$$(a'') \quad \lim_{l \rightarrow \infty} \bar{\theta}(l, T) = \lim_{l \rightarrow \infty} \sup \theta(l, T),$$

$$(b'') \quad \lim_{l \rightarrow \infty} \sup \{\theta(l, T) / \bar{\theta}(l, T)\} = 1.$$

[For the proof see Appendix 1.]

LEMMA 3. Let $b_l^{(0)}(T) \neq 0$ for all l . (1)

Then, if we take a positive integer l^* such that

$$2l^*r_0/V^{1/3} \leq 1 \quad (2)$$

$$\text{and} \quad \bar{\theta}(l^*, T) \cdot 8l^*r_0/V^{1/3} \leq \alpha < 1, \quad (3)$$

then we have for every $l \leq l^*$

$$(i) \quad 0 < 1 - \alpha \leq b_l^\dagger(V, T) / b_l^{(0)}(T) \leq 1 + \alpha < 2,$$

$$(ii) \quad |(1/l) \ln \{b_l^\dagger(V, T) / b_l^{(0)}(T)\}| \leq \bar{\theta}(l^*, T) \cdot 8r_0/V^{1/3} (1 - \alpha).$$

⟨Proof⟩—By (2) we have

$$2lr_0/V^{1/3} \leq 1 \quad \text{for every } l \leq l^*. \quad (4)$$

By (ii) of Lemma 1 we have

$$\theta(l, T) \leq \bar{\theta}(l^*, T) \quad \text{for every } l \leq l^*. \quad (5)$$

Hence by (3)

$$\theta(l, T) \cdot 8lr_0/V^{1/3} \leq \alpha < 1 \quad \text{for every } l \leq l^*. \quad (5)$$

Applying Lemma 2 in view of (4) and (6), we obtain (i) and then (ii), by (5).

§ 1.2. Discrimination between large and small clusters (in the real and the ideal system)

LEMMA 4. For any value of T such that

$$\text{g.l.b.}_{l \geq 1} |b_l^{(0)}(T)|^{1/l} > 0, \quad (1)$$

and for any assigned value of $v(>0)$, we can find a function $l^*(N)$ of $N(=1, 2, 3, \dots)$ which satisfies the following five conditions simultaneously.

1° $l^*(N)$ assumes a positive integral value for every value of N .

2° $(2r_0/v^{1/3})l^*(N)/N^{1/3} \leq 1$ for all values of N except at most a finite number of values of N .

3° There exists a constant $\alpha(0 < \alpha < 1)$ such that $(8r_0/v^{1/3})l^*(N)\bar{\theta}(l^*(N), T)/N^{1/3} \leq \alpha$ for all values of N except at most a finite number of values of N .

4° There exists a number $K(>0)$, independent of N , such that $(\ln N)/l^*(N) \leq K$ for all values of N .

$$5^\circ \quad \lim_{N \rightarrow \infty} \{\bar{\theta}(l^*(N), T)/N^{1/3}\} = 0.$$

«Proof»—Case 1: The case in which the sequence $\{\theta(l, T)\}_{l=1}^{\infty}$ is bounded. In this case the sequence $\{\bar{\theta}(l, T)\}_{l=1}^{\infty}$ is bounded, by (iii) of Lemma γ . Hence 5° holds spontaneously, and we can make 1°, 2°, 3° and 4° hold by choosing a suitable function $l^*(N)$; for example, take

$$l^*(N) = [N^\rho] + 1, \quad (2)^\ddagger$$

where ρ is a constant independent of N such that $0 < \rho < 1/3$. (2a)

Case 2: The case in which the sequence $\{\theta(l, T)\}_{l=1}^{\infty}$ is unbounded. In this case the sequence $\{\bar{\theta}(l, T)\}_{l=1}^{\infty}$ is unbounded, by (iii) of Lemma γ . Let us prove the lemma for this case.

[[1]] By (ii) of Lemma β we have

$$A^{(1)}(l, T) = \text{l.u.b.}_{1 \leq x_l \leq l r_0} |b_l^{(1)}(x_l, T)| < +\infty, \quad (3a)$$

$$A^{(2)}(l, T) = \text{l.u.b.}_{\substack{1 \leq x_l \leq l r_0 \\ 1 \leq y_l \leq l r_0}} |b_l^{(2)}(x_l, y_l, T)| < +\infty, \quad (3b)$$

$$A^{(3)}(l, T) = \text{l.u.b.}_{\substack{0 \leq x_l \leq l r_0 \\ 0 \leq y_l \leq l r_0 \\ 1 \leq z_l \leq l r_0}} |b_l^{(3)}(x_l, y_l, z_l, T)| < +\infty. \quad (3c)$$

Hence by (ii) of Postulate 2

$$(A^{(1)}(l, T))^{1/l} = \text{l.u.b.}_{1 \leq x_l \leq l r_0} |b_l^{(1)}(x_l, T)|^{1/l} \leq G^{(1)}, \quad (4a)$$

similarly

$$(A^{(2)}(l, T))^{1/l} \leq G^{(2)}, \quad (A^{(3)}(l, T))^{1/l} \leq G^{(3)}, \quad (4b, c)$$

where $G^{(1)}$, $G^{(2)}$ and $G^{(3)}$ are positive constants independent of l . Now from (3) of Lemma 1 we have

$$\begin{aligned} 0 \leq \theta(l, T) &\leq \max \{A^{(1)}(l, T), A^{(2)}(l, T), A^{(3)}(l, T)\} + |b_l^{(0)}(T)| \\ &\leq 2 \max \{A^{(1)}(l, T), A^{(2)}(l, T), A^{(3)}(l, T), |b_l^{(0)}(T)|\}, \end{aligned} \quad (5)$$

since by (3a, b, c) of Lemma 1

$$0 \leq \theta^{(i)}(l, T) \leq A^{(i)}(l, T) + |b_l^{(0)}(T)| \quad (i=1, 2, 3). \quad (6)$$

From (5) and (4a, b, c) we have by (i) of Postulate 2

$$\begin{aligned} 0 \leq (\theta(l, T))^{1/l} &\leq 2^{1/l} \max \{(A^{(1)}(l, T))^{1/l}, (A^{(2)}(l, T))^{1/l}, \\ &\quad (A^{(3)}(l, T))^{1/l}, |b_l^{(0)}(T)|^{1/l}\} \\ &\leq 2 \max \{G^{(1)}, G^{(2)}, G^{(3)}, G^{(0)}\} = G \end{aligned} \quad (7)$$

‡ Throughout this paper, $[x]$ means the greatest integer not exceeding x , for any real number x .

say; where G is a positive constant independent of l . That is, the sequence $\{(\theta(l, T))^{1/l}\}_{l=1}^{\infty}$ is bounded. Now by (1) we have

$$|b_l^{(0)}(T)|^{1/l} \geq g^{(0)} > 0, \quad (8)$$

where $g^{(0)}$ is a positive constant independent of l . From (4) of Lemma 2 we have by (7) and (8)

$$0 \leq (\theta(l, T))^{1/l} \leq G/g^{(0)} = G_0 \quad (9)$$

say; where G_0 is a positive constant independent of l . That is, the sequence $\{(\theta(l, T))^{1/l}\}_{l=1}^{\infty}$ is bounded. And here

$$G_0 > 1. \quad (10)$$

For, if we suppose that $G_0 \leq 1$, we have by (9)

$$0 \leq \theta(l, T) \leq G_0^l \leq 1, \quad (11)$$

contrary to the assumption that the sequence $\{\theta(l, T)\}_{l=1}^{\infty}$ is unbounded.

[[2]] By Definition 2, for every l we can find at least one positive integer l' such that $l' \leq l$ and $\theta(l', T) = \bar{\theta}(l', T)$. Hence, for every l such that $\bar{\theta}(l, T) > 1$, we can find at least one positive integer l' such that

$$(\bar{\theta}(l, T))^{1/l} \leq (\bar{\theta}(l, T))^{1/l'} = (\theta(l', T))^{1/l'}. \quad (12)$$

And, for every l such that $\bar{\theta}(l, T) \leq 1$, we have $(\bar{\theta}(l, T))^{1/l} \leq 1$. (13)

From (12), (13), (9) and (10) we have for all l

$$0 \leq (\bar{\theta}(l, T))^{1/l} \leq \max\{G_0, 1\} = G_0 (> 1). \quad (14)$$

That is, the sequence $\{(\bar{\theta}(l, T))^{1/l}\}_{l=1}^{\infty}$ is bounded.

[[3]] If σ and m are positive constants (independent of N) and

$$0 < \sigma < m/(\ln G_0), \quad (15)$$

then we have

$$\lim_{N \rightarrow \infty} \{\bar{\theta}([\sigma \ln N] + 1, T)/N^m\} = 0. \quad (16)$$

For by (14), (10) and (15) we have

$$0 \leq \bar{\theta}([\sigma \ln N] + 1, T)/N^m \leq G_0^{\sigma \ln N + 1}/N^m \leq G_0 N^{\sigma \ln G_0 - m} \rightarrow 0 \quad (N \rightarrow \infty). \quad (17)$$

[[4]] Now put

$$l^*(N) = [\sigma \ln N] + 1, \quad (18)$$

where σ is a constant (independent of N) such that

$$0 < \sigma < 1/(3 \ln G_0). \quad (19)$$

Then the condition 2° , as well as 1° , is satisfied. For we can find an integer $n_0 (> 0)$ such that

$$([\sigma \ln N] + 1) / N^{1/3} \leq v^{1/3} / 2r_0 \quad (N \geq n_0), \quad (20)$$

since the left-hand side tends to zero as $N \rightarrow \infty$. The condition 3° is satisfied. For, for σ so chosen that (19) holds, taking a real number μ such that

$$0 < \sigma < \mu / (\ln G_0) < 1 / (3 \ln G_0), \quad (21)$$

and then, putting $m = \mu$ in (15) and (16), we have

$$\lim_{N \rightarrow \infty} \{\bar{\theta}([\sigma \ln N] + 1, T) / N^\mu\} = 0, \quad (22)$$

and we have

$$\lim_{N \rightarrow \infty} \{([\sigma \ln N] + 1) / N^{1/3-\mu}\} = 0. \quad (23)$$

Hence we can find an integer $n_1 (> 0)$ such that

$$([\sigma \ln N] + 1) \bar{\theta}([\sigma \ln N] + 1, T) / N^{1/3} \leq \alpha v^{1/3} / 8r_0 \quad (N \geq n_1), \quad (24)$$

since the left-hand side tends to zero as $N \rightarrow \infty$, by (22) and (23). The condition 4° is satisfied. For

$$(\ln N) / ([\sigma \ln N] + 1) < 1 / \sigma \quad (\text{for all } N). \quad (25)$$

Finally the condition 5° is satisfied. For, putting $m = 1/3$ in (15) and (16), we have from (19)

$$\lim_{N \rightarrow \infty} \{\bar{\theta}([\sigma \ln N] + 1, T) / N^{1/3}\} = 0. \quad (26)$$

DEFINITION 3. For any value of T for which (1) of Lemma 4 holds and for any value of $v (> 0)$, we take any one of those functions which satisfy the five conditions stated in Lemma 4, and we denote it by $l^*(N)$.

Remark.—In Lemma 4 we have seen that, given T and v , there exist one or more (in general, many) $l^*(N)$'s which satisfy the five conditions. Then we specify arbitrarily one $l^*(N)$ out of the set of all such $l^*(N)$'s for given T and v , and denote this arbitrarily specified $l^*(N)$ by $l^*(N)$ (the same symbol as before). In all the following arguments we shall use this arbitrarily specified $l^*(N)$ for each T and v .

DEFINITION 4. By the "real system" we mean the system which consists of N particles confined within a cubic vessel of volume V , and whose temperature is conserved at $T^\circ K$ by contact with a thermostat, and which therefore follows the canonical distribution law, and whose total potential energy U_N is expressed as the sum of pair potentials $u(r_{ij})$:

$$U_N = \sum_{N \geq i > j \geq 1} u(r_{ij}). \quad (1)$$

We put $V = Nv$, so that v is the volume per particle or the specific volume. Now, for any given $T (> 0)$ and $v (> 0)$, we define the "real system" for every positive integer N , that is to say, we define an infinite sequence of "real systems" ($N = 1, 2, 3, \dots$ to ∞). According to statistical mechanics the configuration integral (i.e., the configurational part of the partition function) of the "real system" is given by†

† The value of the configuration integral $\Omega_N(Nv, T)$ depends not only on the variables N, v and T but also on the shape of the vessel in which all the N particles are confined. But here we have specified the shape to be cubic. Then we denote this specification by the superscript †. On the contrary, the (0)-system never concerns any domain or vessel.

$$\mathcal{Q}_N^{\dagger}(Nv, T) = \frac{1}{N!} \int_V \exp\left(-\frac{U_N}{kT}\right) d\{N\}. \quad (2)$$

Then according to the Ursell development this is expressed as

$$\mathcal{Q}_N^{\dagger}(Nv, T) = \sum_{m_l \geq 0} \left(\sum_{l=1}^N m_l = N \right) \prod_{l=1}^N \frac{\{Nv b_l^{\dagger}(Nv, T)\}^{m_l}}{m_l!}, \quad (3)$$

where the sum extends over all the sets $\{m_l\}_{l=1}^N$ of non-negative integers m_l which satisfy the condition stated in the brackets, and the $b_l^{\dagger}(Nv, T)$'s are the real cluster integrals given in Definition 1. We may also say that we have defined the "real system" as the system whose configuration integral is given by (3); this definition seems more convenient to us.

On the other hand, for any given $T(>0)$ and $v(>0)$, we define the "(0)-system"—for every positive integer N —as the system whose configuration integral $\mathcal{Q}_N^{(0)}$ is given by the following expression, that is to say, we define an infinite sequence of "(0)-systems" ($N=1, 2, 3, \dots$ to ∞).

$$\mathcal{Q}_N^{(0)}(Nv, T) = \sum_{m_l \geq 0} \left(\sum_{l=1}^N m_l = N \right) \prod_{l=1}^N \frac{\{Nv b_l^{(0)}(T)\}^{m_l}}{m_l!}, \quad (4)$$

where the $b_l^{(0)}(T)$'s are the non-boundary cluster integrals given in Definition 1.

Now from 2° of Lemma 4 we have

$$\lim_{N \rightarrow \infty} \{l^*(N)/N\} = 0. \quad (5)$$

Hence we can find an integer $\nu_0(>0)$ such that

$$l^*(N) < N \quad (N \geq \nu_0). \quad (6)$$

Hence, for all $N \geq \nu_0$, we may express (3) and (4) as follows.

First, by

$$\tilde{T}_N^{\dagger}(i|Nv, T) = \prod_{l=1}^{l^*(N)} \frac{\{Nv b_l^{\dagger}(Nv, T)\}^{m_l}}{m_l!} \quad (7)$$

we denote any one of those terms of the sum on the right-hand side of (3) in which $m_l=0$ for every l such that $l^*(N)+1 \leq l \leq N$; where the index i on the left-hand side represents the set $\{m_l\}_{l=1}^N = \{m_1, m_2, \dots, m_{l^*(N)}, 0, 0, \dots, 0\}$ given on the right-hand side. Second, by

$$\tilde{T}_N^{\dagger}(j|Nv, T) = \prod_{l=1}^{l^*(N)} \frac{\{Nv b_l^{\dagger}(Nv, T)\}^{m_l}}{m_l!} \prod_{l=l^*(N)+1}^N \frac{\{Nv b_l^{\dagger}(Nv, T)\}^{m_l}}{m_l!} \quad (8)$$

we denote any one of those terms of the sum on the right-hand side of (3) in which $m_l \geq 1$ for at least one value of l such that $l^*(N)+1 \leq l \leq N$; where the index j on the left-hand side represents the set $\{m_l\}_{l=1}^N$ given on the right-hand side. Then (3)

may be written[‡]

$$\Omega_N^\dagger(Nv, T) = \phi_N^\dagger(Nv, T) + \Psi_N^\dagger(Nv, T) \quad (9)$$

where

$$\phi_N^\dagger(Nv, T) = \sum_i \tilde{T}_N^\dagger(i|Nv, T) \quad (9a)$$

$$\Psi_N^\dagger(Nv, T) = \sum_j \tilde{\tilde{T}}_N^\dagger(j|Nv, T). \quad (9b)$$

Similarly, for the "(0)-system", we write

$$\tilde{T}_N^{(0)}(i|Nv, T) = \prod_{l=1}^{P(N)} \frac{\{Nv b_l^{(0)}(T)\}^{m_l}}{m_l!}, \quad (10)$$

where the index i represents the same set of m_l that the same index i in (7) represents. And we write

$$\tilde{\tilde{T}}_N^{(0)}(j|Nv, T) = \prod_{l=1}^{P(N)} \frac{\{Nv b_l^{(0)}(T)\}^{m_l}}{m_l!} \prod_{l=P(N)+1}^N \frac{\{Nv b_l^{(0)}(T)\}^{m_l}}{m_l!}, \quad (11)$$

where the index j represents the same set of m_l that the same index j in (8) represents. Then (4) may be written

$$\Omega_N^{(0)}(Nv, T) = \phi_N^{(0)}(Nv, T) + \Psi_N^{(0)}(Nv, T) \quad (12)$$

where

$$\phi_N^{(0)}(Nv, T) = \sum_i \tilde{T}_N^{(0)}(i|Nv, T) \quad (12a)$$

$$\Psi_N^{(0)}(Nv, T) = \sum_j \tilde{\tilde{T}}_N^{(0)}(j|Nv, T). \quad (12b)$$

§ 1.3. Relation between the small cluster parts of the real and the ideal system

LEMMA 5. For any value of T such that (1) of Lemma 4 holds and

$$b_l^{(0)}(T) > 0 \quad \text{for every } l, \quad (1)$$

and for any assigned value of $v(>0)$,

(i) $\tilde{T}_N^\dagger(i|Nv, T) > 0$ for each i , and hence $\phi_N^\dagger(Nv, T) > 0$, for every value of N except at most a finite number of values of N ,

$$(ii) \lim_{N \rightarrow \infty} (1/N) \{\ln \phi_N^\dagger(Nv, T) - \ln \phi_N^{(0)}(Nv, T)\} = 0.$$

⟨Proof⟩—[[1]] From (1) we have for every N

$$\tilde{T}_N^{(0)}(i|Nv, T) > 0, \quad \tilde{\tilde{T}}_N^{(0)}(j|Nv, T) > 0 \quad \text{for each } i \text{ and each } j. \quad (2)$$

[‡] A similar technique is used in a theory of condensing systems⁽⁷⁾, in which we attempt to derive the liquid phase by overcoming a difficulty of the Mayer theory of condensation.

Now, in view of the conditions 2° and 3° (of Lemma 4) satisfied by $l^*(N)$, we can apply Lemma 3, putting $V=Nv$. Then, for almost † every N , we have

$$0 < 1 - \alpha \leq b_i^\dagger(Nv, T)/b_i^{(0)}(T) \leq 1 + \alpha < 2 \quad (l \leq l^*(N)), \quad (3)$$

$$\text{whence by (1)} \quad b_i^\dagger(Nv, T) > 0 \quad (l \leq l^*(N)). \quad (4)$$

Hence we have for almost every N

$$\tilde{T}_N^\dagger(i|Nv, T) > 0 \quad \text{for each } i. \quad (5)$$

Hence (i). And then we have for almost every N

$$\left| \frac{1}{l} \ln \frac{b_i^\dagger(Nv, T)}{b_i^{(0)}(T)} \right| \leq \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)} \quad (l \leq l^*(N)). \quad (6)$$

Hence for almost every N we have for each i

$$\begin{aligned} \left| \frac{1}{N} \ln \tilde{T}_N^\dagger(i|Nv, T) - \frac{1}{N} \ln \tilde{T}_N^{(0)}(i|Nv, T) \right| &\leq \sum_{l=1}^{P_N(N)} \frac{m_l}{N} \left| \ln \frac{b_i^\dagger(Nv, T)}{b_i^{(0)}(T)} \right| \\ &\leq \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)}, \end{aligned} \quad (7)$$

since

$$\sum_{l=1}^{P_N(N)} l m_l = N \quad \text{for each } i. \quad (8)$$

[[2]] Now by $i=m^\dagger$ we mean the set of m_l for the greatest term of the sum $\sum_i \tilde{T}_N^\dagger(i|Nv, T) [= \phi_N^\dagger(Nv, T)]$, and by $i=m^{(0)}$ we mean the set of m_l for the greatest term of the sum $\sum_i \tilde{T}_N^{(0)}(i|Nv, T) [= \phi_N^{(0)}(Nv, T)]$, and we let \tilde{P}_N be the total number of terms of the sum $\sum_i \tilde{T}_N^\dagger(i|Nv, T)$ —and, therefore, of the sum $\sum_i \tilde{T}_N^{(0)}(i|Nv, T)$. Then by (5) we have for almost every N

$$\tilde{T}_N^\dagger(m^\dagger|Nv, T) \leq \phi_N^\dagger(Nv, T) \leq \tilde{P}_N \tilde{T}_N^\dagger(m^\dagger|Nv, T), \quad (9a)$$

and by (2) we have for every N

$$\tilde{T}_N^{(0)}(m^{(0)}|Nv, T) \leq \phi_N^{(0)}(Nv, T) \leq \tilde{P}_N \tilde{T}_N^{(0)}(m^{(0)}|Nv, T), \quad (9b)$$

and we have

$$\tilde{T}_N^\dagger(m^\dagger|Nv, T) \geq \tilde{T}_N^\dagger(m^{(0)}|Nv, T) \quad (10a)$$

$$\tilde{T}_N^{(0)}(m^\dagger|Nv, T) \leq \tilde{T}_N^{(0)}(m^{(0)}|Nv, T). \quad (10b)$$

Now let P_N be the total number of terms of the sum $\sum (\sum l m_l = N) \dots$ in (3) and (4) of Definition 4. Then

$$\tilde{P}_N < P_N. \quad (11)$$

† "Almost every N " means "every N except at most a finite number of N ".

And P_N is the *partitio numerorum* of N ; we have asymptotically

$$\ln P_N \sim \pi (2N/3)^{1/2} \quad (\text{as } N \rightarrow \infty). \quad (12)$$

Hence, putting $\xi(N) = (1/N) \ln \tilde{P}_N$, (13)

we have $\lim_{N \rightarrow \infty} \xi(N) = 0$. (14)

From (9a) and (13) we have for almost every N

$$0 \leq \frac{1}{N} \ln \phi_N^+(Nv, T) - \frac{1}{N} \ln \tilde{T}_N(m^+ | Nv, T) \leq \xi(N). \quad (15a)$$

From (9b) and (13) we have for every N

$$0 \leq \frac{1}{N} \ln \phi_N^{(0)}(Nv, T) - \frac{1}{N} \ln \tilde{T}_N^{(0)}(m^{(0)} | Nv, T) \leq \xi(N). \quad (15b)$$

Now, applying (7) to the terms for which $i = m^+, m^{(0)}$, we have for almost every N

$$\left| \frac{1}{N} \ln \tilde{T}_N^+(m^+ | Nv, T) - \frac{1}{N} \ln \tilde{T}_N^{(0)}(m^+ | Nv, T) \right| \leq \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)} \quad (16a)$$

$$\left| \frac{1}{N} \ln \tilde{T}_N^+(m^{(0)} | Nv, T) - \frac{1}{N} \ln \tilde{T}_N^{(0)}(m^{(0)} | Nv, T) \right| \leq \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)}. \quad (16b)$$

From (10a, b) and (16a, b) we have[†] for almost every N .

$$\left| \frac{1}{N} \ln \tilde{T}_N^+(m^+ | Nv, T) - \frac{1}{N} \ln \tilde{T}_N^{(0)}(m^{(0)} | Nv, T) \right| \leq \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)}. \quad (17)$$

From (15a, b) and (17) we have^{††} for almost every N

$$\left| \frac{1}{N} \ln \phi_N^+(Nv, T) - \frac{1}{N} \ln \phi_N^{(0)}(Nv, T) \right| \leq \xi(N) + \frac{\bar{\theta}(l^*(N), T)}{N^{1/3}} \cdot \frac{8r_0}{v^{1/3}(1-\alpha)}, \quad (18)$$

which proves (ii), by (14) and by the condition 5° (of Lemma 4) satisfied by $l^*(N)$.

§ 1.4. Some properties of the ideal system

Argument 1. (On the “(0)-system.”) Here we argue about the “(0)-system”, i.e., the system for which the volume dependence of the cluster integrals is completely neglected. We may say that such a system is a sort of an *ideal system*.^{††}

[[1]] For any given T , let $\chi_T(\zeta)$ be the analytic function (of ζ) defined by the following power series, whose radius of convergence is $R(T)$ [$0 < R(T) \leq +\infty$], and by its analytical continuation.

[†] If $a_1 \geq a_2$, $b_1 \leq b_2$, $|a_1 - b_1| \leq \eta$, $|a_2 - b_2| \leq \eta$, then $-\eta \leq a_2 - b_2 \leq a_1 - b_1 \leq \eta$.

^{††} If $0 \leq A - a_1 \leq \xi$, $0 \leq B - b_2 \leq \xi$, then $|(A - a_1) - (B - b_2)| \leq \xi$. Hence $|A - B| \leq |(A - a_1) - (B - b_2)|$

+ $|a_1 - b_2| \leq \xi + \eta$ if $|a_1 - b_2| \leq \eta$.

^{†††} And here we admit the use of the ideal numbers $+\infty$ and $-\infty$.

$$\chi_T(\zeta) = \sum_{l=1}^{\infty} l b_l^{(0)}(T) \zeta^l. \quad (1)$$

And, for any given T and v , let $z(v, T)$ be the least of those real positive values of ζ which satisfy

$$\text{either (a) } \chi_T(\zeta) = 1/v \text{ or (b) } \chi_T(\zeta) \text{ is singular.} \quad (2)$$

Then it has been proved by Kahn and Uhlenbeck⁽⁶⁾ and by Born and Fuchs⁽⁵⁾ that, for every value of v ($0 < v < +\infty$), we have

$$\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^{(0)}(v, T) = \sum_{l=1}^{\infty} v b_l^{(0)}(T) (z(v, T))^l - \ln z(v, T). \quad (3)$$

[[2]] For any given T , let $z_s(T)$ be the singularity of $\chi_T(\zeta)$ nearest to the origin on the real positive axis of ζ ; $0 < z_s(T) \leq +\infty$.

$$\text{Then we have} \quad R(T) \leq z_s(T). \quad (4)$$

$$\text{We put} \quad \chi_T(z_s(T)) = 1/v_s(T), \quad (5)$$

where we admit even the cases that $\pm \infty = 1/0$ and that $0 = 1/(\pm \infty)$.

Assume that the analytic function $\chi_T(\zeta)$ has a positive differential coefficient everywhere in the real interval $0 \leq \zeta \leq z_s(T)$. This assumption is necessary and sufficient in order that the function $z(v, T)$ of $1/v$, in which we agree $z(v, T) = 0$ for $1/v = 0$, should have neither an infinite differential coefficient nor a discontinuity anywhere in the interval $0 \leq 1/v < +\infty$. On this assumption, we first get

$$\chi_T(\zeta) > 0 \quad (0 < \zeta \leq z_s(T)) \quad (6)$$

$$+\infty \geq \chi_T(z_s(T)) > 0 \quad \text{i.e.} \quad 0 \leq v_s(T) < +\infty, \quad (7)$$

and then get the following results (i) and (ii).

(i) Through the interval (of v) $v_s(T) < v < +\infty$, $z(v, T)$ is determined by (a) of (2), and throughout this interval of v the function $z(v, T)$ of v is regular and positive and has a negative differential coefficient, and $\lim[v \rightarrow v_s(T)] z(v, T) = z_s(T)$, $\lim[v \rightarrow +\infty] z(v, T) = 0$; hence, throughout this interval of v , the function $\omega^{(0)}(v, T)$ of v , i.e., (3) is, regular and has a positive differential coefficient, and the function of v

$$\dot{\omega}^{(0)}(v, T) = (\partial/\partial v) \omega^{(0)}(v, T) = \sum_{l=1}^{\infty} l b_l^{(0)}(T) (z(v, T))^{l-1} \quad (8)$$

is regular and positive and has a negative differential coefficient, and $\lim[v \rightarrow +\infty] \dot{\omega}^{(0)}(v, T) = 0$. For, putting $z = z(v, T)$, so $v = v(z, T)$, and regarding z as an argument, then we have

$$\dot{\omega}^{(0)}(v(z, T), T) = \sum_{l=1}^{\infty} l b_l^{(0)}(T) z^{l-1}, \quad \lim_{z \rightarrow 0} \dot{\omega}^{(0)}(v(z, T), T) = 0, \quad (9)$$

$$(\partial/\partial z) \dot{\omega}^{(0)}(v(z, T), T) = \sum_{l=1}^{\infty} l b_l^{(0)}(T) z^{l-2} = \chi_T(z)/z > 0 \quad (0 < z < z_s(T)), \quad (10)$$

and hence throughout the interval $0 < z < z_s(T)$ the function $\dot{\omega}^{(0)}(v(z, T), T)$ of z is regular and positive and has a positive differential coefficient.

(ii) Through the interval (of v) $0 < v \leq v_s(T)$, $z(v, T)$ is determined by (b) of (2) and is independent of v , and

$$z(v, T) = z_s(T) \quad (0 < v \leq v_s(T)), \quad (11)$$

and hence $\omega^{(0)}(v, T)$, (3), is a linear function of v , and so $\dot{\omega}^{(0)}(v, T)$ is independent of v .

In order that the statement of (ii) should not be meaningless, it is necessary and sufficient that

$$0 < z_s(T) < +\infty, \quad 0 < v_s(T) < +\infty. \quad (12)$$

In this case the function $\omega^{(0)}(v, T)$ of v —as also $\dot{\omega}^{(0)}(v, T)$ and $z(v, T)$ —consists of two parts which are analytically independent of each other, and each of which is regular over its own interval— $(0, v_s(T))$, $(v_s(T), +\infty)$ —, the boundary point $v = v_s(T)$ being irregular.

[[3]] It is considered that, for all sufficiently low temperatures, that is, for all the temperatures lower than some temperature which may be a little lower than the critical temperature, we have

$$b_l^{(0)}(T) > 0 \quad \text{for every } l. \quad (I)$$

For every value of T such that (I) holds,

(i) the radius of convergence, $R(T)$, of the power series (1) is given by

$$R(T) = z_s(T), \quad (13)$$

(ii) the derivative of the power series (1) is positive everywhere in the interval $0 \leq \zeta \leq R(T) = z_s(T)$, and hence (6), (7) and (i), (ii) of [[2]] are all true.

[[4]] Mayer and others' arguments show that, whenever we are at a temperature T for which (I) holds,

$$\lim_{l \rightarrow \infty} (b_l^{(0)}(T))^{1/l} = b_0^{(0)}(T) \quad \text{exists and} \quad 0 < b_0^{(0)}(T) < +\infty. \quad (II)$$

If this holds, we have by Cauchy-Hadamard's theorem

$$b_0^{(0)}(T) = 1/R(T) = \limsup_{l \rightarrow \infty} |b_l^{(0)}(T)|^{1/l}. \quad (14)$$

From this and (13) we have

$$1/b_0^{(0)}(T) = z_s(T). \quad (15)$$

[[5]] Now see Lemma 6, which concerns the "(0)-system" only. In this lemma we first assume (I) [there denoted by (1)] for T and then assume (II) [denoted by (2)] as an accepted property of $b_l^{(0)}(T)$ and bear in mind a proved fact [denoted by (5) \rightarrow (6) \rightarrow (8)] which was discussed in [[1]] and [[2]]. And what we have to prove is the truth of (9).

[[6]] Also in Theorems α and β we assume—though not explicitly stated there—(I) [i.e., (1) of Lemma 6] for $b_l^{(0)}(T)$ (together with (II) [i.e., (2) of Lemma 6]) and bear in mind the proved fact (5) \rightarrow (6) \rightarrow (8) (of Lemma 6). That both (i) and

(II) are true is sufficient in order that both (i) of Postulate 2 and (1) of Lemma 4 may be true. For, if a sequence $\{a_n\}_{n=1}^{\infty}$ ($a_n > 0$ for all n) converges to λ ($0 < \lambda < +\infty$), then, taking ϵ ($0 < \epsilon < \lambda$), we can find an integer m (> 0) such that

$$0 < \lambda - \epsilon < a_n < \lambda + \epsilon \quad (n \geq m). \quad (16)$$

Hence we have for all n

$$0 < \min\{a_1, a_2, \dots, a_{m-1}, \lambda - \epsilon\} \leq a_n \leq \max\{a_1, a_2, \dots, a_{m-1}, \lambda + \epsilon\}. \quad (17)$$

Thus, in proving Theorems α and β , we can use the so far given lemmas as well as Lemma 6.

Lemma δ . For any value of T such that (1) of Lemma 4 holds and for any assigned value of v (> 0), we can find a function $l'(N)$ of N ($= 1, 2, 3, \dots$) which satisfies the following four conditions simultaneously.

1° $l'(N)$ assumes a positive integral value for every N .

2° $l'(N) \leq l^*(N)$ for every N .

3° There exists an integer $N'(\epsilon)$ (> 0), corresponding to any $\epsilon > 0$, such that

$$l'(N)/N^{1/2} < \epsilon \quad \text{for all } N \geq N'(\epsilon).$$

4° There exists a number K' (> 0), independent of N , such that

$$(\ln N)/l'(N) \leq K' \quad \text{for all } N.$$

[For, by the condition 4° (of Lemma 4) satisfied by $l^*(N)$, it is proved that we can find a function $l'(N)$ which satisfies the conditions 2° and 4° of Lemma δ . Hence it is obvious that we can find $l'(N)$ which satisfies 1°, 2°, 3° and 4°.]

LEMMA 6. Let us be at a temperature T such that

$$b_l^{(0)}(T) > 0 \quad \text{for every } l, \quad (1)$$

let

$$\lim_{l \rightarrow \infty} (b_l^{(0)}(T))^{1/l} = b_0^{(0)}(T) \text{ exist and } 0 < b_0^{(0)}(T) < +\infty, \quad (2)$$

and put

$$z_s(T) = 1/b_0^{(0)}(T), \quad (3)$$

$$v_s(T) = \left\{ \sum_{l=1}^{\infty} l b_l^{(0)}(T) (z_s(T))^l \right\}^{-1}. \quad (4)$$

Then, if

$$v > v_s(T) \quad (5)$$

and hence

$$z(v, T) < z_s(T), \quad (6)$$

where $z(v, T)$ is the least positive real root of the equation

$$\sum_{l=1}^{\infty} l v b_l^{(0)}(T) z^l = 1, \quad (7)$$

i.e., if we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln Q_N^{(0)}(Nv, T) = \omega^{(0)}(v, T) = \sum_{l=1}^{\infty} v b_l^{(0)}(T) (z(v, T))^l - \ln z(v, T). \quad (8)$$

then we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \phi_N^{(0)}(Nv, T) = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Omega_N^{(0)}(Nv, T) [= \omega^{(0)}(v, T)]. \quad (9)$$

《Proof》—Now, assuming T to be fixed, we shall omit the parameter T , so that we shall write $b_i^{(0)}$, z_s , v_s for $b_i^{(0)}(T)$, $z_s(T)$, $v_s(T)$ respectively; and moreover we shall write z for $z(v, T)$.

[[1]] For any value of $v (> v_s)$ and for any value of $N (\geq \nu_0)$ we choose a term $\tilde{T}_N^{(0)}(\hat{i}|Nv)$ out of $\sum_i \tilde{T}_N^{(0)}(i|Nv)$ in the following way. Take

$$\tilde{T}_N^{(0)}(\hat{i}|Nv) = \prod_{l=1}^{l^*(N)} \frac{(Nv b_l^{(0)})^{\hat{m}_l}}{\hat{m}_l!} \quad (10)$$

$$\text{where} \quad \hat{m}_l = [\bar{m}_l] + \mu_l \quad \text{for} \quad 1 \leq l \leq l'(N), \quad (11a)$$

$$\hat{m}_l = 0 \quad \text{for} \quad l'(N) + 1 \leq l \leq l^*(N). \quad (11b)$$

Here $l'(N)$ is a function given by Lemma δ ; and (11b) is omitted if $l'(N) = l^*(N)$. And here

$$\bar{m}_l = Nv b_l^{(0)} z^l \quad \text{for all } l, \quad (12)$$

where z is the least positive real root (which exists if $v > v_s$) of (7). And

$$\mu_l = [\mu(N)] + \nu_l \quad \text{for} \quad 1 \leq l \leq l'(N), \quad (13)$$

where

$$\mu(N) = \frac{2}{l'(N)(l'(N)+1)} \cdot (N - \sum_{l=1}^{l'(N)} l[\bar{m}_l]) \quad (14)$$

and $\nu_l (l=1, 2, \dots, l'(N))$ are such that:

$$1^\circ \text{ each } \nu_l \text{ is either } 0 \text{ or } 1, \quad (15)$$

$$2^\circ \sum_{l=1}^{l'(N)} l\nu_l = (1/2)l'(N)(l'(N)+1) \{\mu(N) - [\mu(N)]\}. \quad (16)$$

[Here obviously the right-hand side of (16) is non-negative and less than $(1/2)l'(N) \cdot (l'(N)+1)$, and is proved to be an integer, by (14). Hence by Appendix 2 we can find such a set of ν_l .]

Now from (1) and (12) we have

$$\bar{m}_l > 0 \quad \text{for all } l. \quad (17)$$

And from (7) and (12) we have

$$\sum_{l=1}^{\infty} l\bar{m}_l = N. \quad (18)$$

From this and (17) we have

$$N - \sum_{l=1}^{l'(N)} l[\bar{m}_l] > 0 \quad \text{i.e.} \quad \mu(N) > 0, \quad (19)$$

by (14). Hence by (13) and (15)

$$\mu_l \geq 0 \quad \text{for} \quad 1 \leq l \leq l'(N). \quad (20)$$

Hence by (11a, b) and (17)

$$\hat{m}_l \geq 0 \quad \text{for} \quad 1 \leq l \leq l^*(N). \quad (21)$$

Now from (11a, b), (13), (14) and (16) we have

$$\sum_{l=1}^{l^*(N)} l \hat{m}_l = N. \quad (22)$$

We have thus proved that the term given by (10) is really a term of the sum \sum_i .
 [[2]] Hereafter, \sum' means summation from $l=1$ to $l=l'(N)$, and $\sum'(C_1, C_2, \dots, C_n)$ means summation over those l for which $1 \leq l \leq l'(N)$ and the conditions C_1, C_2, \dots, C_n are satisfied. And \sum'' means $\sum'(m_l \geq 1)$, and $\sum''(C_1, C_2, \dots, C_n)$ means $\sum'(m_l \geq 1, C_1, C_2, \dots, C_n)$.

Now, according to Stirling's formula, we have

$$\ln \hat{m}_l! = \hat{m}_l (\ln \hat{m}_l - 1) + (1/2) \ln \hat{m}_l + \gamma(\hat{m}_l) \quad (23)$$

for every positive integer \hat{m}_l , where $\gamma(\hat{m}_l)$ is bounded, that is, there is an absolute constant $K_1 (> 0)$ such that

$$|\gamma(\hat{m}_l)| \leq K_1 \quad \text{for every} \quad \hat{m}_l \geq 1. \quad (24)$$

From (10) we have by (11b) and (23)

$$(1/N) \ln \tilde{T}_N^{(0)}(i|Nv) = (1/N) \sum' \{ \hat{m}_l \ln (Nv b_l^{(0)}) - \hat{m}_l \ln \hat{m}_l + \hat{m}_l - (1/2) \ln \hat{m}_l - \gamma(\hat{m}_l) \}. \quad (25)$$

From this and (8) we have

$$(1/N) \ln \tilde{T}_N^{(0)}(i|Nv) - \omega^{(0)}(v) = A(N) + B(N) + C(N) + D(N), \quad (26)$$

where

$$A(N) = (1/N) \sum' \hat{m}_l \{ \ln (Nv b_l^{(0)}) - \ln \hat{m}_l \} + \ln z, \quad (26a)$$

$$B(N) = (1/N) \sum' \hat{m}_l - \sum_{l=1}^{\infty} v b_l^{(0)} z_l, \quad (26b)$$

$$C(N) = (1/N) \sum'' (-1/2) \ln \hat{m}_l, \quad (26c)$$

$$D(N) = (1/N) \sum'' (-\gamma(\hat{m}_l)). \quad (26d)$$

With the use of (11a) and (12), (26a) becomes

$$A(N) = A_1(N) + A_2(N) + A_3(N) + A_4(N) + A_5(N), \quad (27)$$

where

$$A_1(N) = (1/N) \sum' l (\bar{m}_l - \hat{m}_l) \ln z, \quad (27a)$$

$$A_2(N) = (1/N) \sum'' (\bar{m}_l - \hat{m}_l) \ln \{ ([\bar{m}_l] + \mu_l) / \bar{m}_l \}, \quad (27b)$$

$$A_3(N) = -(1/N) \sum' \bar{m}_l \ln \{ ([\bar{m}_l] + \mu_l) / \bar{m}_l \}, \quad (27c)$$

$$A_4(N) = \ln z - (1/N) \sum' \bar{m}_l \ln z, \quad (27d)$$

$$A_5(N) = (1/N) \sum' (\hat{m}_l = 0) \bar{m}_l \ln z. \quad (27e)$$

Then we have (cf. Appendix 3)

$$|A_1(N)| \leq (1/N) \sum'' l |\bar{m}_l - \hat{m}_l| |\ln z| < (1/N) \sum' l (\mu_l + 1) |\ln z|, \quad (28)^3$$

and

$$|A_2(N)| \leq (1/N) \sum'' |\bar{m}_l - \hat{m}_l| |\ln \{ ([\bar{m}_l] + \mu_l) / \bar{m}_l \}| < P(N) + Q(N), \quad (29)$$

where

$$\begin{aligned} 0 \leq P(N) &= (1/N) \sum'' (\bar{m}_l \geq 1) (\mu_l + 1) \ln (\mu_l + 2) \\ &\leq (1/N) l'(N) (\mu(N) + 2) \ln (N + 3) \leq (1/N) (l'(N))^2 (\mu(N) + 2) K'' \end{aligned} \quad (29a)^{3,A;7,8}$$

(where K'' is a positive constant independent of N)*, and

$$\begin{aligned} 0 \leq Q(N) &= (1/N) \sum'' (\bar{m}_l < 1) \mu_l \ln (\mu_l / \bar{m}_l) \\ &= (1/N) \sum'' (\bar{m}_l < 1) \mu_l |\ln \{ N v f(l) (z/z_s)^l / \mu_l \}| \end{aligned} \quad (29b)^{2,4}$$

where**

$$\lim_{l \rightarrow \infty} (1/l) \ln f(l) = 0. \quad (29b_1)$$

Then

$$Q(N) = Q_1(N) + Q_2(N) + Q_3(N) + Q_4(N), \quad (30)$$

where $0 \leq Q_1(N) = (1/N) \sum'' (\bar{m}_l < 1) \mu_l \ln (N / \mu_l)$

$$\leq (1/N) l'(N) (\mu(N) + 1) \ln N \leq (1/N) (l'(N))^2 (\mu(N) + 1) K' \quad (30a)^{7,8}$$

(where K' is a positive constant independent of N) (see 4° of Lemma δ), and

$$0 \leq Q_2(N) = (1/N) \sum'' (\bar{m}_l < 1) \mu_l |\ln v| \leq (1/N) \sum' \mu_l |\ln v|, \quad (30b)$$

$$0 \leq Q_3(N) = (1/N) \sum'' (\bar{m}_l < 1) \mu_l |\ln f(l)| \leq (1/N) \sum_{l=1}^{h-1} \mu_l |\ln f(l)| + (1/N) \sum' l \mu_l \quad (30c)^{***}$$

* By 4° of Lemma δ we can find $K''(>0)$ such that $\ln(N+2)/l'(N) \leq K''$ for all N .

** By (2) we have

$$b_l^{(0)}(T) = f(l, T) (b_0^{(0)}(T))^l \quad (2')$$

where

$$\lim_{l \rightarrow \infty} (1/l) \ln f(l, T) = 0. \quad (2'')$$

In (29b) we employ this $f(l, T)$, written as $f(l)$, and use (3) and (12).

$$0 \leq Q_3(N) \leq (1/N) \sum_{l=1}^{h-1} \mu_l |\ln f(l)| + (1/N) \sum_{l=h}^{l'(N)} l \mu_l \quad (\text{if } l'(N) \geq h), \quad (30c_1)$$

$$0 \leq Q_3(N) \leq (1/N) \sum_{l=1}^{h-1} \mu_l |\ln f(l)| \quad (\text{if } l'(N) < h). \quad (30c_2)$$

(where b is a positive integer independent of N and v , such that $|\ln f(l)| \leq l$ whenever $l \geq b$; such b exists by (29b₁)), and

$$0 \leq Q_4(N) = (1/N) \sum'' (\bar{m}_l < 1) l \mu_l |\ln(z/z_s)| \leq (1/N) \sum' l \mu_l |\ln(z/z_s)|. \quad (30d)$$

Next

$$|A_3(N)| \leq (1/N) \sum'' |m_l \ln \{([\bar{m}_l] + \mu_l) / \bar{m}_l\}| < X(N) + Y(N), \quad (31)$$

where

$$0 \leq X(N) = (1/N) \sum'' (\mu_l = 0) K_0 \leq K_0 l'(N) / N, \quad (31a)^{5,B}$$

$$0 \leq Y(N) = (1/N) \sum'' (\mu_l \geq 1) \mu_l \leq (1/N) \sum' \mu_l, \quad (31b)^C$$

(where K_0 is a positive absolute constant). Next

$$|A_4(N)| = \sum_{l=l'(N)+1}^{\infty} l v b_l^{(0)} z^l |\ln z|, \quad (32)$$

$$|A_5(N)| \leq (1/N) \sum' (\hat{m}_l = 0) l \bar{m}_l |\ln z| < (1/N) \sum' l |\ln z|. \quad (33)^6$$

Now (26b) becomes

$$B(N) = B_1(N) + B_2(N), \quad (34)$$

where

$$B_1(N) = - \sum_{l=l'(N)+1}^{\infty} v b_l^{(0)} z^l, \quad B_2(N) = (1/N) \sum' (\hat{m}_l - \bar{m}_l). \quad (34a,b)$$

Then

$$|B_2(N)| \leq (1/N) \sum' |\hat{m}_l - \bar{m}_l| < (1/N) \sum' \mu_l + l'(N) / N. \quad (35)^3$$

From (26c, d)

$$|C(N)| = (1/N) \sum'' (1/2) \ln \hat{m}_l \leq (1/2N) l'(N) \ln N, \quad (36)^9$$

$$|D(N)| = (1/N) \sum'' |\gamma(\hat{m}_l)| \leq K_1 l'(N) / N \quad (37)$$

(where K_1 is a positive absolute constant) (see (24)).

[[3]] By 3° of Lemma δ , there are integers $N''(\epsilon) (> 0)$ and $N'''(\epsilon) (> 0)$, corresponding to any $\epsilon > 0$, such that

$$|0 < l'(N) / N \leq (1/N) \sum' l| = (1/2N) l'(N) (l'(N) + 1) < \epsilon \text{ for } N \geq N''(\epsilon), \quad (38)$$

$$|0 < (1/N) l'(N) \ln N| < \epsilon \quad \text{for } N \geq N'''(\epsilon). \quad (39)$$

By 4° of Lemma δ , there is an integer $N_0(M) (> 0)$, corresponding to any $M > 0$, such that

$$l'(N) \geq M \quad \text{for } N \geq N_0(M). \quad (40)$$

Now if $v \sim v_s(z/z_s)$, the series $\sum_{l=1}^{\infty} v b_l^{(0)} z^l$ and $\sum_{l=1}^{\infty} l v b_l^{(0)} z^l (=1)$ are convergent, that is, there are integers $M_0(\epsilon) (> 0)$ and $M_1(\epsilon) (> 0)$, corresponding to any $\epsilon > 0$, such that

$$|0 < \sum_{l=l'+1}^{\infty} v b_l^{(0)} z^l| < \epsilon \quad \text{for } l' \geq M_0(\epsilon), \quad (41)$$

$$[0 <] \sum_{l=l'(N)+1}^{\infty} l v b_l^{(0)} z^l < \epsilon \quad \text{for } l' \geq M_1(\epsilon). \quad (42)$$

Hence, putting $M=M_0(\epsilon)$, $M_1(\epsilon)$, in (40), we have, for any $\epsilon > 0$,

$$[0 <] \sum_{l=l'(N)+1}^{\infty} v b_l^{(0)} z^l < \epsilon \quad \text{for } N \geq N_0(M_0(\epsilon)), \quad (43)$$

$$[0 <] \sum_{l=l'(N)+1}^{\infty} l v b_l^{(0)} z^l < \epsilon \quad \text{for } N \geq N_0(M_1(\epsilon)). \quad (44)$$

Now from (13), (16) and (14) we have

$$0 < \sum' l \mu_l = (1/2) l'(N) (l'(N) + 1) \mu'(N) = N - \sum' l [\bar{m}_l]. \quad (45)$$

Hence by (18) and (12)

$$0 < (1/N) \sum' l \mu_l = (1/N) \{ \sum' l (\bar{m}_l - [\bar{m}_l]) + \sum_{l=l'(N)+1}^{\infty} l \bar{m}_l \} < (1/N) \sum' l + \sum_{l=l'(N)+1}^{\infty} l v b_l^{(0)} z^l. \quad (46)$$

Hence by (38) and (44) we have for any $\epsilon > 0$

$$[0 < (1/N) \sum' l \mu_l \leq] (1/N) \sum' l \mu_l < \epsilon \quad \text{for } N \geq N_1(\epsilon), \quad (47)$$

where

$$N_1(\epsilon) = \max \{ N''(\epsilon/2), N_0(M_1(\epsilon/2)) \}. \quad (47a)$$

Now from (45), (46) and (47), we have, for any $\epsilon > 0$,

$$[0 <] (1/N) (l'(N))^2 (\mu(N) + 1) [< (4/N) \sum' l \mu_l] < \epsilon \quad \text{for } N \geq N_1(\epsilon/4), \quad (48a)$$

$$[0 <] (1/N) (l'(N))^2 (\mu(N) + 2) [< (6/N) \sum' l \mu_l] < \epsilon \quad \text{for } N \geq N_1(\epsilon/6). \quad (48b)$$

[[4]] From [[2]] and [[3]], we have, for any $\epsilon > 0$,

- 1) $|A_1(N)| < \epsilon$ for $N \geq N^{(1)}(\epsilon) = \max \{ N_1(\epsilon/2|\ln z|), N''(\epsilon/2|\ln z|) \},$
- 2) $0 \leq P(N) < \epsilon$ for $N \geq N^{(2)}(\epsilon) = N_1(\epsilon/6K''),$
- 3) $0 \leq Q_1(N) < \epsilon$ for $N \geq N^{(3)}(\epsilon) = N_1(\epsilon/4K'),$
- 4) $0 \leq Q_2(N) < \epsilon$ for $N \geq N^{(4)}(\epsilon) = N_1(\epsilon/|\ln v|),$
- 5) $0 \leq Q_3(N) < \epsilon$ for $N \geq N^{(5)}(\epsilon) = \max \{ (2/\epsilon) \sum_{l=1}^{h-1} \mu_l |\ln f(l)|, N_1(\epsilon/2) \},$
- 6) $0 \leq Q_4(N) < \epsilon$ for $N \geq N^{(6)}(\epsilon) = N_1(\epsilon/|\ln(z/z_s)|),$
- 7) $0 \leq X(N) < \epsilon$ for $N \geq N^{(7)}(\epsilon) = N''(\epsilon/K_0),$
- 8) $0 \leq Y(N) < \epsilon$ for $N \geq N^{(8)}(\epsilon) = N_1(\epsilon),$
- 9) $|A_2(N)| < \epsilon$ for $N \geq N^{(9)}(\epsilon) = N_0(M_1(\epsilon/|\ln z|)),$
- 10) $|A_3(N)| < \epsilon$ for $N \geq N^{(10)}(\epsilon) = N''(\epsilon/|\ln z|),$
- 11) $|B_1(N)| < \epsilon$ for $N \geq N^{(11)}(\epsilon) = N_0(M_0(\epsilon)),$
- 12) $|B_2(N)| < \epsilon$ for $N \geq N^{(12)}(\epsilon) = \max \{ N_1(\epsilon/2), N''(\epsilon/2) \},$
- 13) $|C(N)| < \epsilon$ for $N \geq N^{(13)}(\epsilon) = N''(2\epsilon),$
- 14) $|D(N)| < \epsilon$ for $N \geq N^{(14)}(\epsilon) = N''(\epsilon/K_1).$

(49)

Therefore, putting

$$N_2(\epsilon) = \max \{N^{(1)}(\epsilon), N^{(2)}(\epsilon), N^{(3)}(\epsilon), \dots, N^{(14)}(\epsilon)\}, \quad (50)$$

we can, for any given $\epsilon > 0$, choose such a term $\tilde{T}_N^{(0)}(\hat{i}|Nv)$ out of $\sum_i \tilde{T}_N^{(0)}(i|Nv)$ for every $N \geq \max \{\nu_0, N_2(\epsilon/28)\}$, that

$$|(1/N) \ln \tilde{T}_N^{(0)}(\hat{i}|Nv) - \omega^{(0)}(v)| \leq |A_1(N)| + P(N) + Q_1(N) + \dots + |D(N)| < \epsilon/2. \quad (51)$$

Now by (8) there is an integer $N_3(\epsilon) (> 0)$, corresponding to $\epsilon (> 0)$, such that

$$|(1/N) \ln \mathcal{Q}_N^{(0)}(Nv) - \omega^{(0)}(v)| < \epsilon/2 \quad \text{for } N \geq N_3(\epsilon). \quad (52)$$

By (51) and (52), we can, for any $\epsilon > 0$, choose such a term $\tilde{T}_N^{(0)}(\hat{i}|Nv)$ out of $\sum_i \tilde{T}_N^{(0)}(i|Nv)$ for every $N \geq N_4(\epsilon) = \max \{\nu_0, N_2(\epsilon/28), N_3(\epsilon)\}$, that

$$|(1/N) \ln \tilde{T}_N^{(0)}(\hat{i}|Nv) - (1/N) \ln \mathcal{Q}_N^{(0)}(Nv)| < \epsilon. \quad (53)$$

Now from (1) we have (2) of Lemma 5 and hence we have for every N

$$(1/N) \ln \tilde{T}_N^{(0)}(\hat{i}|Nv) \leq (1/N) \ln \phi_N^{(0)}(Nv) \leq (1/N) \ln \mathcal{Q}_N^{(0)}(Nv). \quad (54)$$

From (53) and (54) we have for any $\epsilon > 0$

$$|(1/N) \ln \phi_N^{(0)}(Nv) - (1/N) \ln \mathcal{Q}_N^{(0)}(Nv)| < \epsilon \quad \text{for } N \geq N_4(\epsilon). \quad (55)$$

This proves (9).

Argument 2 (On the function $l^*(N)$.) The conditions 2° , 3° and 5° of Lemma 4 makes the increase of the function $l^*(N)$ (with N) *slow* enough to make Lemma 5 hold. And the condition 4° of Lemma 4 makes the increase of the function $l^*(N)$ (with N) *rapid* enough to make Lemma 6 hold. In fact, in the proof of Lemma 5 we have used 1° , 2° , 3° and 5° of Lemma 4, but not 4° of Lemma 4. And in the proof of Lemma 6 (essentially, of Lemma $\hat{\alpha}$) we have used 1° and 4° of Lemma 4, but neither 2° nor 3° nor 5° of Lemma 4. Therefore, if we regard Lemma 5 as independent of any other lemma, then, in place of $l^*(N)$ in Lemma 5, we should use $l_r^*(N)$, for which 1° , 2° , 3° and 5° of Lemma 4 hold. And, if we regard Lemma 6 as independent of any other lemma, then, in place of $l^*(N)$ in Lemma 6 (and consequently in Lemma $\hat{\alpha}$), we should use $l_r^*(N)$, for which 1° and 4° of Lemma 4 hold and

$$l_r^*(N) \leq N \quad \text{for almost every } N. \quad (1)$$

[In Lemmas $\hat{\alpha}$ and 6 with $l^*(N)$ replaced by $l_r^*(N)$, we write $l_r'(N)$ in place of $l'(N)$.] Then the most trivial of $l_r^*(N)$ is $l_r^*(N) = \text{constant}$, and the most trivial of $l_r^*(N)$ is $l_r^*(N) = N$. Now if we denote by $\{f(N)\}$ the set of the values (for any given N) of all those functions $f(N)$ which satisfy the specified conditions, then we have for almost every N

$$\{l_s^*(N)\} \cap \{l_r^*(N)\} = \{l^*(N)\} \supseteq \{l'(N)\} \quad (2)$$

$$\{l_r^*(N)\} \supset \{l'_r(N)\} \supset \{l^*(N)\} \subset \{l_s^*(N)\}. \quad (3)$$

Argument 3 (On the physical meaning of Lemma 6.) As is stated in Argument 2, we now regard Lemma 6 as independent of any other lemma or theorem. Then Lemma 6 has by itself a meaning, which is as follows. In the proof of Lemma 6 we have picked out a term (denoted by \hat{i}) by which the whole configuration integral of the "(0)-system" in gaseous state for sufficiently large N is expressed with as small error as we please, and by which therefore the whole configuration integral for infinitely large N is rigorously expressed; in a word, this is a predominant term. And we have found that the distribution of sizes* of clusters for such a term is given by \hat{m}_i of (11a, b) and the largest size of clusters that exists in such a term should be of the order of at least $\ln N$ [by 4° of Lemma δ] and at most $N^{1/2-\delta}$ (where $\delta(>0)$ may be as small as we please) [by 3° of Lemma δ]. Hence, for a macroscopic "(0)-system" ($N \sim 10^{23}$) in gaseous state, it is practically sufficient to take the first $\ln N$ (~ 10 to 100) terms in the cluster series and it is practically impossible to find any cluster of the size of the order of or greater than $N^{1/2} (\sim 10^{10}$ to $10^{12})$.

§ 1.5. Theorem on equality between the thermodynamic functions of the real and the ideal system

THEOREM α . For any value of T such that $b_i^{(0)}(T) > 0$ for every i , and for any value of v such that $v > v_s(T)$;—

[I] The necessary and sufficient condition that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln Q_N^{\dagger}(Nv, T) = \omega^{\dagger}(Nv, T) \text{ should exist} \quad (\alpha)$$

and $\omega^{\dagger}(v, T) = \omega^{(0)}(v, T)$

is that:

$$\lim_{N \in E_{(+)} } \sup_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^{\dagger}(Nv, T) - \ln \phi_N^{\dagger}(Nv, T) \} \leq 0 \quad (A)^{**}$$

when the set $E_{(+)} = E\{N(\text{positive integers}); \Psi_N^{\dagger}(Nv, T) > 0\}$ is an infinite set, and

$$\ln \left(1 + \frac{\Psi_N^{\dagger}(Nv, T)}{\phi_N^{\dagger}(Nv, T)} \right)^{-1} = o(N) \quad \text{for } N \in E_{(-)} \quad (B)$$

when the set $E_{(-)} = E\{N(\text{positive integers}); \Psi_N^{\dagger}(Nv, T) \leq 0\}$ is an infinite set.

(Note) If $E_{(+)}$ [or $E_{(-)}$] is a finite or empty set, then the condition (A) [or (B)] should be omitted. Of course, always $E_{(+)} + E_{(-)} = E\{1, 2, 3, \dots\}$: an enumerable infinite set.

* The size of a cluster means the number of molecules composing the cluster, that is, l .

** $\lim_{N \in E, N \rightarrow \infty} \sup_{a_N}$ means the greatest limit of the sequence, $\{a_N\}_{N \in E}$, formed by the terms for $N \in E$ of the sequence $\{a_N\} (N=1 \dots \infty)$. And $E\{x; C\}$ means the set of all x that satisfy the condition C .

[I'] The theorem [I] is still true if the condition (A) is replaced by

$$\lim_{N \in E_{(+)}'} \sup_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^{\dagger}(Nv, T) - \ln \phi_N^{(0)}(Nv, T) \} \leq 0. \quad (A')$$

[II] When the set $E_{(+)}$ is an infinite set, we have

$$\begin{aligned} & \lim_{N \in E_{(+)}'} \sup_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^{\dagger}(Nv, T) - \ln \phi_N^{\dagger}(Nv, T) \} \\ &= \lim_{N \in E_{(+)}'} \sup_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^{\dagger}(Nv, T) - \ln \phi_N^{(0)}(Nv, T) \} \\ &\leq \lim_{N \in E_{(+)}'} \sup_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^{\dagger}(Nv, T) - \ln \Psi_N^{(0)}(Nv, T) \}. \end{aligned}$$

Remark 1.—Here, from the purely mathematical viewpoint, we have not assumed even the existence of the limit $\omega^{\dagger}(v, T)$ and so we should obtain the condition for existence and equality, the invalidity of which implies either inequality or nonexistence; but, from the physical viewpoint, we require the existence of this limit, from which the thermodynamic quantities (free energy, pressure, etc.) should be derived.

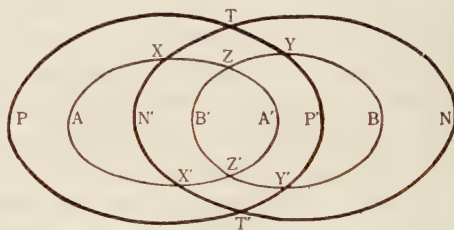


Fig. 1.1.

Remark 2.—In the theorem [I], the domains of validity of the conditions (α) , (A), (B), (P) $\equiv \{E_{(+)} \text{ is an infinite set}\}$, and (N) $\equiv \{E_{(-)} \text{ is an infinite set}\}$ are illustrated by Fig. 1.1, in which (P) = PTP'T' [which means that (P) is true inside the closed curve PTP'T' and untrue outside it], (N) = NTN'T', (A) = AXA'X', (B) = BYB'Y', then $(\alpha) = AXN'X' + BYP'Y' + A'ZB'Z'$.

(1)

[Proof of [I]]—Hereafter, for brevity's sake we write ϕ_N , Ψ_N for $\phi_N(Nv, T)$, $\Psi_N(Nv, T)$ regarded as functions of N .

Case 1: The case in which $E_{(-)}$ is at most a finite set, that is,

$$\Psi_N^{\dagger} > 0 \quad \text{for almost every } N. \quad (2)$$

[1] Proof of the sufficiency of (A). First let inequality in (A) hold, that is, let

$$\lim_{N \rightarrow \infty} \sup (1/N) \ln (\Psi_N^{\dagger} / \phi_N^{\dagger}) = \bar{\lambda}, \quad \text{where } -\infty \leq \bar{\lambda} < 0. \quad (3)$$

Then [by (i) and (iii) of I of Appendix 4] we have for almost every N

$$(1/N) \ln (\Psi_N^{\dagger} / \phi_N^{\dagger}) < 0 \quad \text{i.e.} \quad \Psi_N^{\dagger} / \phi_N^{\dagger} < 1. \quad (4)$$

Hence [by (2) and Lemma 5 (i)] we have for almost every N

$$0 < (1/N) \ln (1 + \Psi_N^{\dagger} / \phi_N^{\dagger}) < (1/N) \ln 2. \quad (5)$$

Hence

$$\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) = 0. \quad (6)$$

Next let equality in (A) hold, that is, let

$$\limsup_{N \rightarrow \infty} (1/N) \ln(\Psi_N^\dagger / \phi_N^\dagger) = 0. \quad (7)$$

Then [by (i) of I of Appendix 4] for any $\epsilon > 0$ we have for almost every N

$$(1/N) \ln(\Psi_N^\dagger / \phi_N^\dagger) < \epsilon. \quad (8)$$

Hence, for every N (except at most a finite number of N) such that $\Psi_N^\dagger / \phi_N^\dagger \geq 1$, we have

$$0 < (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) \leq (1/N) \ln 2 + (1/N) \ln(\Psi_N^\dagger / \phi_N^\dagger) < (1/N) \ln 2 + \epsilon. \quad (9)$$

And, for every N (except at most a finite number of N) such that $\Psi_N^\dagger / \phi_N^\dagger < 1$, we have [by (2) and Lemma 5 (i)]

$$0 < (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) < (1/N) \ln 2. \quad (10)$$

By (9) and (10), for any $\epsilon > 0$ we have $0 < (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) < 2\epsilon$ for almost every N . Hence

$$\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) = 0. \quad (11)$$

Now by (2) of Definition 4 we have

$$\mathcal{Q}_N^\dagger(Nv, T) > 0 \quad \text{for every } N. \quad (12)$$

By this and Lemma 5(i), Eq. (9) of Definition 4 may be written

$$(1/N) \ln \mathcal{Q}_N^\dagger(Nv, T) = (1/N) \ln \phi_N^\dagger + (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) \quad (13)$$

for almost every N . By Lemma 5 (ii) and Lemma 6, we have for $v > v_s(T)$

$$\lim_{N \rightarrow \infty} (1/N) \ln \phi_N^\dagger = \lim_{N \rightarrow \infty} (1/N) \ln \phi_N^{(0)} = \omega^{(0)}(v, T). \quad (14)$$

From (6), (11), (13) and (14) [by Corol. 3 of III of Appendix 4] we obtain (α).

[[2]] Proof of the necessity of (A). Let (A) be untrue, that is, let

$$\limsup_{N \rightarrow \infty} (1/N) \ln(\Psi_N^\dagger / \phi_N^\dagger) = \bar{\lambda}, \quad \text{where } 0 < \bar{\lambda} \leq +\infty. \quad (15)$$

Then, since [by (2) and Lemma 5 (i)] we have for almost every N

$$(1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) > (1/N) \ln(\Psi_N^\dagger / \phi_N^\dagger), \quad (16)$$

we have [by II of Appendix 4]

$$+\infty \geq \limsup_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) \geq \bar{\lambda}, \quad \text{where } 0 < \bar{\lambda} \leq +\infty. \quad (17)$$

Hence the sequence $\{(1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger)\}_{N=1}^\infty$ has at least one *limit-value** (finite or

* A "limit-value" means an accumulating value or *Häufungspunkt*.

infinite) greater than zero. Hence $\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger)$ either exists as a positive value (finite or $+\infty$) or is nonexistent. Hence, by (13) and (14), $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^\dagger(Nv, T)$ either exists as a value (finite or $+\infty$) greater than $\omega^{(0)}(v, T)$ or is nonexistent, [by Corol. 2 and 3 of III of Appendix 4].

This completes the proof of [I] for Case 1.

Case 2: The case in which $E_{(+)}$ is at most a finite set, that is,

$$\Psi_N^\dagger \leq 0 \quad \text{for almost every } N. \quad (18)$$

[[1]] Proof of the sufficiency of (B). Let (B) be true, that is, let

$$\lim_{N \rightarrow \infty} (-1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger) = 0. \quad (19)$$

Then by (13) and (14) [by Corol. 3 of III of Appendix 4] we obtain (α) .

[[2]] Proof of the necessity of (B). Let (B) be untrue. Then $\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^\dagger / \phi_N^\dagger)$ is either nonexistent or exists as a negative value (finite or $-\infty$) [by (18) and Lemma 5 (i)]. Hence by (13) and (14) [by Corol. 2 and 3 of III of Appendix 4] $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^\dagger(Nv, T)$ is either nonexistent or exists as a value (finite or $-\infty$) less than $\omega^{(0)}(v, T)$.

This completes the proof of [I] for Case 2.

Case 3: The case in which both $E_{(+)}$ and $E_{(-)}$ are infinite sets, that is, there are an infinite number of N such that $\Psi_N^\dagger > 0$ and there are an infinite number of N such that $\Psi_N^\dagger \leq 0$.

Using the increasing sequence $\{N_v\}_{v=1}^\infty$ of all $N \in E_{(+)}$ and the increasing sequence $\{N'_v\}_{v=1}^\infty$ of all $N \in E_{(-)}$, we divide the sequence $\{(1/N) \ln \mathcal{Q}_N^\dagger(Nv, T)\}_{N=1}^\infty (s)$ into two partial sequences, viz. $\{(1/N_v) \ln \mathcal{Q}_{N_v}^\dagger(N_v v, T)\}_{v=1}^\infty (s_1)$ and $\{(1/N'_v) \ln \mathcal{Q}_{N'_v}^\dagger(N'_v v, T)\}_{v=1}^\infty (s_2)$.

[[1]] Let both (A) and (B) be true. Then, first, from the truth of (A), by the same reasoning for the partial sequence (s_1) as that for the sequence (s) in [[1]] of Case 1, we obtain

$$\lim_{v \rightarrow \infty} (1/N_v) \ln \mathcal{Q}_{N_v}^\dagger(N_v v, T) = \omega^{(0)}(v, T). \quad (20)$$

Next, from the truth of (B), by the same reasoning for the partial sequence (s_2) as that for the sequence (s) in [[1]] of Case 2, we obtain

$$\lim_{v \rightarrow \infty} (1/N'_v) \ln \mathcal{Q}_{N'_v}^\dagger(N'_v v, T) = \omega^{(0)}(v, T). \quad (21)$$

From (20) and (21) we obtain (α) .

[[2]] Let at least one of (A) and (B) be untrue. Then, by applying the reasoning for the sequence (s) in [[2]] of Cases 1 and 2 to the partial sequences (s_1) and (s_2) , we find that at least one of $\lim_{v \rightarrow \infty} (1/N_v) \ln \mathcal{Q}_{N_v}^\dagger(N_v v, T)$ and $\lim_{v \rightarrow \infty} (1/N'_v) \ln \mathcal{Q}_{N'_v}^\dagger(N'_v v, T)$ is either nonexistent or exists as a value (finite or infinite) different from $\omega^{(0)}(v, T)$. Especially, if both exist, they cannot have the same value. Hence $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^\dagger(Nv, T)$ does not exist in any case.

By [[1]] and [[2]] it is proved that for Case 3 the truth of both (A) and (B) is necessary and sufficient for the truth of (α) . Thus the proof of [I] is completed.

⟨Proof of [I']⟩—For almost all N that belong to the infinite set $E_{(+)}$, we have [by (i) and (2) of Lemma 5]

$$(1/N) (\ln \Psi_N^+ - \ln \phi_N^+) = (1/N) (\ln \Psi_N^+ - \ln \phi_N^{(0)}) + (1/N) (\ln \phi_N^{(+)} - \ln \phi_N^+). \quad (22)$$

Now by (ii) of Lemma 5

$$\lim_{N \rightarrow \infty} (1/N) (\ln \phi_N^{(+)} - \ln \phi_N^+) = 0. \quad (23)$$

From (22) and (23) we obtain [by Corol. 1 of III of Appendix 4]

$$\lim_{N \in E_{(+)}} \sup_{N \rightarrow \infty} (1/N) (\ln \Psi_N^+ - \ln \phi_N^+) = \lim_{N \in E_{(+)}} \sup_{N \rightarrow \infty} (1/N) (\ln \Psi_N^+ - \ln \phi_N^{(0)}). \quad (24)$$

Hence the conditions (A) and (A') are equivalent. This proves [I].

⟨Proof of [II]⟩—The first equality in [II] is Eq. (24), just proved. We shall then prove the rest. For all N that belong to the infinite set $E_{(+)}$, we have [by (2) of Lemma 5]

$$\begin{aligned} (1/N) (\ln \Psi_N^+ - \ln \phi_N^{(0)}) &= (1/N) (\ln \Psi_N^+ - \ln \Psi_N^{(0)}) + (1/N) (\ln \Psi_N^{(+)} - \ln \phi_N^{(0)}) \\ &< (1/N) (\ln \Psi_N^+ - \ln \Psi_N^{(0)}) + (1/N) \ln(1 + \Psi_N^{(+)} / \phi_N^{(0)}). \end{aligned} \quad (25)$$

Now [by (2) of Lemma 5] Eq. (12) of Definition 4 may be written

$$(1/N) \ln Q_N^{(0)}(Nv, T) = (1/N) \ln \phi_N^{(0)} + (1/N) \ln(1 + \Psi_N^{(0)} / \phi_N^{(0)}) \quad (26)$$

for every N . But by Lemma 6 we have for $v > v_s(T)$

$$\lim_{N \rightarrow \infty} \{ (1/N) \ln Q_N^{(0)}(Nv, T) - (1/N) \ln \phi_N^{(0)} \} = 0. \quad (27)$$

Hence by (26) we have

$$\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^{(0)} / \phi_N^{(0)}) = 0. \quad (28)$$

Finally from (25) [by II of Appendix 4] and from (28) [by Corol. 1 of III of Appendix 4] we obtain

$$\lim_{N \in E_{(+)}} \sup_{N \rightarrow \infty} (1/N) (\ln \Psi_N^+ - \ln \phi_N^{(+)}) \leq \lim_{N \in E_{(+)}} \sup_{N \rightarrow \infty} (1/N) (\ln \Psi_N^+ - \ln \Psi_N^{(0)}). \quad (29)$$

Corollary 1. If, in the theorem [I], we replace (A) by

$$\lim_{N \in E_{(+)}} \sup_{N \rightarrow \infty} (1/N) (\ln \Psi_N^+ - \ln \Psi_N^{(0)}) \leq 0 \quad (A_1)$$

and leave (B) as it is, then we have a sufficient condition for the truth of (α) .

⟨Proof⟩—Use [I] and [II] of the theorem.

Corollary 2. For any value of T such that $b_l^{(+)}(T) > 0$ for every l , and for any value of v such that $v > v_s(T)$, a sufficient condition that

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \ln Q_N^l(Nv, T) = \omega^l(v, T) \quad \text{should exist} \\ \text{and} \quad \omega^l(v, T) = \omega^{(0)}(v, T) \end{aligned} \right\} \quad (\alpha)$$

is that, for every N except at most a finite number of N ,

$$0 \leq b_l^l(Nv, T) \leq b_l^{(0)}(T) \quad \text{provided that } l \leq N. \quad (\text{C})$$

⟨Proof⟩—Let (C) be true. Then by (8) and (11) of Definition 4 we have for each j

$$0 \leq \tilde{T}_N^+(j|Nv, T) \leq \tilde{T}_N^{(0)}(j|Nv, T) \quad (1)$$

for every N except at most a finite number of N . Now if the set $E_{(+)}$ is an infinite set, then, for the sequence $\{(1/N)(\ln \mathcal{P}_N^+ - \ln \mathcal{P}_N^{(0)})\}_{N \in E_{(+)}}$, the following condition is satisfied except for at most a finite number of terms, because of (1).

$$(1/N)(\ln \mathcal{P}_N^+ - \ln \mathcal{P}_N^{(0)}) \leq 0. \quad (2)$$

Hence this sequence cannot have any limit-values (finite or infinite) greater than zero. That is, the greatest limit does not exceed zero. Therefore the condition (A_1) of Corollary 1 is satisfied.

Now by (1) we have for each j

$$\tilde{T}_N^+(j|Nv, T) = 0 \quad (3)$$

for every $N \in E_{(-)}$ except for at most a finite number of N . Hence, if the set $E_{(-)}$ is an infinite set, we have

$$\ln(1 + \mathcal{P}_N^+/\phi_N^l)^{-1} = 0 \quad (4)$$

for every $N \in E_{(-)}$ except for at most a finite number of N . [Note that there are only at most finite number of N such that $\phi_N^l = 0$, by Lemma 5 (i).] Therefore the condition (B) of the theorem [I] is satisfied.

Thus by the theorem [I] and Corollary 1, (α) holds. Hence the corollary.

§ 1.6. Theorem on inequality between the thermodynamic functions of the real and the ideal system

THEOREM 1². For any value of T such that $b_l^{(0)}(T) > 0$ for every l , and for any value of v such that $v > v_s(T)$;—

[III] The necessary and sufficient condition that

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \ln Q_N^l(Nv, T) = \omega^l(v, T) \quad \text{should exist} \\ \text{and} \quad \omega^l(v, T) > \omega^{(0)}(v, T) \end{aligned} \right\} \quad (\beta)$$

is that we should have

$$\Psi_N^+(Nv, T) > \phi_N^+(Nv, T) [> 0] \quad (D_1)$$

for every N except at most a finite number of N and

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^+(Nv, T) - \ln \phi_N^+(Nv, T) \} = \lambda_1(v, T) \quad \text{should exist} \\ \text{and} \quad \lambda_1(v, T) > 0. \end{aligned} \right\} \quad (D_2)$$

And, in case the condition is satisfied, we have

$$\omega^+(v, T) = \omega^{(0)}(v, T) + \lambda_1(v, T). \quad (\beta_1)$$

[III'] The theorem [III] is still true if the condition (D_2) is replaced by the following:

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \{ \ln \Psi_N^+(Nv, T) - \ln \phi_N^{(0)}(Nv, T) \} = \lambda_1(v, T) \quad \text{should exist} \\ \text{and} \quad \lambda_1(v, T) > 0. \end{aligned} \right\} \quad (D_2')$$

[IV] The necessary and sufficient condition that

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Omega_N^+(Nv, T) = \omega^+(v, T) \quad \text{should exist} \\ \text{and} \quad \omega^+(v, T) < \omega^{(0)}(v, T) \end{aligned} \right\} \quad (\gamma)$$

is that we should have

$$-\phi_N^+(Nv, T) < \Psi_N^+(Nv, T) < 0 \quad (E_1)$$

for every N except at most a finite number of N and

$$\left. \begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \ln \{ 1 + \Psi_N^+(Nv, T) / \phi_N^+(Nv, T) \} = \lambda_2(v, T) \quad \text{should exist} \\ \text{and} \quad \lambda_2(v, T) < 0. \end{aligned} \right\} \quad (E_2)$$

And, in case the condition is satisfied, we have

$$\omega^+(v, T) = \omega^{(0)}(v, T) + \lambda_2(v, T). \quad (\gamma_1)$$

⟨Proof of [III]⟩—[[1]] Sufficiency. Let (D_1) and (D_2) be true. [Here by Lemma 5 (i) we have $\phi_N^+ > 0$ for almost every N .] Then by (D_1) we have for almost every N

$$\Psi_N^+ / \phi_N^+ > 1. \quad (1)$$

Hence we have for almost every N

$$(1/N) \ln (\Psi_N^+ / \phi_N^+) < (1/N) \ln (1 + \Psi_N^+ / \phi_N^+) < (1/N) \ln 2 + (1/N) \ln (\Psi_N^+ / \phi_N^+). \quad (2)$$

Hence by (D₂)

$$\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^+ / \phi_N^+) = \lambda_1(v, T) > 0. \quad (3)$$

Hence by (13) and (14) of Theorem α we have

$$\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T) = \omega^{(0)}(v, T) + \lambda_1(v, T). \quad (4)$$

Hence (β) and (β_1).

[[2]] Necessity. By (12) of Theorem α , we have

$$\Psi_N^+ + \phi_N^+ > 0 \text{ or } 1 + \Psi_N^+ / \phi_N^+ > 0 \text{ (for almost every } N). \quad (5)$$

So (1°) if $-\phi_N^+ < \Psi_N^+ \leq 0$ for an infinite number of N , then

$$(1/N) \ln(1 + \Psi_N^+ / \phi_N^+) \leq 0 \text{ (for an infinite number of } N). \quad (6)$$

Hence the sequence $\{(1/N) \ln(1 + \Psi_N^+ / \phi_N^+)\}_{N=1}^{\infty}$ has one or more non-positive limit-values.

Hence by (13) and (14) of Theorem α , $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T)$ is either nonexistent or else $\leq \omega^{(0)}(v, T)$. (2°) If $0 < \Psi_N^+ \leq \phi_N^+$ for an infinite number of N , then

$$0 < (1/N) \ln(1 + \Psi_N^+ / \phi_N^+) \leq (1/N) \ln 2 \text{ (for an infinite number of } N). \quad (7)$$

Hence zero is a limit-value of $\{(1/N) \ln(1 + \Psi_N^+ / \phi_N^+)\}_{N=1}^{\infty}$. Hence by (13) and (14) of Theorem α , $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T)$ is either nonexistent or else $= \omega^{(0)}(v, T)$. (3°)

If (D₁) holds but (D₂) does not hold, then by (2), $\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^+ / \phi_N^+)$ is either nonexistent or else ≤ 0 . Hence by (13) and (14) of Theorem α , $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T)$ is either nonexistent or else $\leq \omega^{(0)}(v, T)$.

<Proof of [III]> By (24) of Theorem α , if and only if (D₂') holds, (D₂) holds.

<Proof of [IV]>—[[1]] Sufficiency. Let (E₁) and (E₂) be true. Then from (E₂) and by (13) and (14) of Theorem α , we have

$$\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T) = \omega^{(0)}(v, T) + \lambda_2(v, T) \quad [\lambda_2(v, T) < 0]. \quad (8)$$

Hence (γ) and (γ_1).

[[2]] Necessity. Let us note (5). (1°) If $\Psi_N^+ \geq 0$ for an infinite number of N , then [by Lemma 5(i)]

$$(1/N) \ln(1 + \Psi_N^+ / \phi_N^+) \geq 0 \text{ (for an infinite number of } N). \quad (9)$$

Hence the sequence $\{(1/N) \ln(1 + \Psi_N^+ / \phi_N^+)\}_{N=1}^{\infty}$ has one or more non-negative limit-values.

Hence by (13) and (14) of Theorem α , $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T)$ is either nonexistent or else $\geq \omega^{(0)}(v, T)$. (2°) If (E₁) holds but (E₂) does not hold, then $\lim_{N \rightarrow \infty} (1/N) \ln(1 + \Psi_N^+ / \phi_N^+)$ is either nonexistent or else ≥ 0 . Hence by (13) and (14) of Theorem

α , $\lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N^+(Nv, T)$ is either non-existent or else $\geq \omega^{(0)}(v, T)$.

Part II

General Discussion of Phase Transitions

§ 2.0. In this article, from the analytical viewpoint we shall give a general discussion (analytical but not quantitative) on the phase transitions (of first order) of real gases, lattice gases, and other systems. And we shall make clear what essential problems we have to solve about the phase transitions.

§ 2.1. For the analytical (or function-theoretical) researches of the phase transitions, we may utilize the following methods of statistical mechanics, which may give the rigorous (though not so concrete) analytical expressions in the limit of infinite system.

Method α . ((Handling of the Ursell development for a canonical ensemble.)) Using the cluster integrals $b_l(V, T)$, we obtain the Ursell development [(5) of § 0.1] of the configuration integral $\mathcal{Q}_N(V, T)$ [(2) of § 0.1] for a real gas. Then the pressure p is given by

$$p/kT = \varphi_{c,N}(V, T) \equiv (\partial/\partial V) \ln \mathcal{Q}_N(V, T). \quad (1)$$

In the limit when $N \rightarrow \infty$, $V = Nv \rightarrow \infty$ with v fixed, we have methods^{(4), (5), (6), (7)} for handling the Ursell development and calculating

$$\omega \equiv \omega(v, T) \equiv \lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N(Nv, T) \quad (2)$$

and

$$\varphi_c \equiv \varphi_c(v, T) \equiv \lim_{N \rightarrow \infty} \varphi_{c,N}(Nv, T), \quad (3)$$

etc., which are directly related to the thermodynamic properties for an infinitely large system. In the argument of the phase transitions, we examine the limiting functions ω and φ_c , etc.

Method β . ((Power-series development for a grand canonical ensemble.)) The grand partition function $\Xi_v(z, T)$ for a real gas is expressed by a power-series of the "fugacity" z of the gas as follows:

$$\Xi_v(z, T) = \sum_{N \geq 0} \mathcal{Q}_N(V, T) z^N. \quad (4)$$

Then the pressure p is given by⁽⁸⁾

$$p/kT = \varphi_{g,v}(z, T) \equiv (1/V) \ln \Xi_v(z, T) = \sum_{l \geq 1} b_l(V, T) z^l, \quad (5)$$

using (5) of § 0.1. [But here the first equality of (5) is only asymptotically valid (by assuming $\ln \Xi_v \propto V$) when V becomes large for real systems. Strictly speaking, we have

$$p/kT = (\partial/\partial V) (V \varphi_{g,v}(z, T)) \equiv (\partial/\partial V) \ln \Xi_v(z, T) \quad (5')$$

for any finite V .] In the argument of the phase transitions, we examine the limiting function

$$\varphi_g \equiv \varphi_g(z, T) \equiv \lim_{V \rightarrow \infty} \varphi_{g,v}(z, T). \quad (6)$$

If one admits that, for real systems, the physical results from the canonical ensemble and from the grand canonical ensemble are the same, then the two functions φ_c and φ_g should be identical. Strictly speaking, we have yet to discuss mathematically the question as to the identity of φ_c and φ_g . But here we will simply denote φ_c and φ_g by φ , unifying them.

Methods α and β are applied not only to real gases but also to lattice gases and crystals.¹²⁾

Method γ . ((Solution of the eigenvalue problem in crystal statistics.)) This method^{10), 10), 11)} is applied to nearest neighbour systems in crystal statistics—a lattice gas, a binary alloy, and an Ising ferromagnet, etc. (for a grand canonical ensemble or a canonical ensemble). Let λ_i ($i=1, 2, \dots, 2^{mn}$) be the eigenvalues of the matrix $|M_{ij}|$ ($i, j=1, 2, \dots, 2^{mn}$) [where $M_{ij}=M_{ij}(z, T)$] defined in the method,^{10), 10), 11)} that is, the roots of the secular equation

$$\det |M_{ij} - \lambda \delta_{ij}| = 0, \quad (7)$$

where the system consists of l lattice-planes, each consisting of $m \times n$ lattice-points, and follows the Born-von Kármán boundary condition. Then the [grand] partition function* is given by

$$\Phi_{l,mn}(z, T) = \sum_{i=1}^{2^{mn}} \lambda_i^l(mn; z, T). \quad (8)$$

Let λ_{\max} be the largest of the λ_i 's. Then we have (for $l \rightarrow \infty$)

$$\bar{\varphi}_{mn}(z, T) \equiv \lim_{l \rightarrow \infty} (1/lmn) \Phi_{l,mn}(z, T) = (1/mn) \ln \lambda_{\max}(mn; z, T). \quad (9)$$

In the argument of the phase transitions, we examine the function $\bar{\varphi}_{mn}$ and its limit when $m \times n \rightarrow \infty$,

$$\bar{\varphi} \equiv \bar{\varphi}(z, T) \equiv \lim_{m \times n \rightarrow \infty} \bar{\varphi}_{mn}(z, T). \quad (10)$$

§ 2.2. Infinite systems and phase transitions.

(Method α) If $N \rightarrow \infty$, $\varphi_{c,N}(Nv, T)$ as well as $\Omega_N(Nv, T)$ may be considered as a regular function of v . Only if $N \rightarrow \infty$, we may expect singularities, which represent phase transitions.

(Method β) Assuming the rigid core (or repulsive force) of the molecules, if $V \rightarrow \infty$, $\Xi_V(z, T)$ is a polynomial in z , so it is regular (and positive since all $\Omega_N > 0$) for all positive z ; hence $\varphi_{g,V}(z, T)$ is regular for all $z > 0$. Only if $V \rightarrow \infty$, we may have singularities.⁸⁾

(Method γ)—If $m \times n \rightarrow \infty$, $|M_{ij}|$ is a finite matrix and all $M_{ij} > 0$, and hence by Frobenius' theorem the largest λ_i is positive and simple everywhere. Only if $m \times n \rightarrow \infty$,

* For a lattice gas, (8)=grand partition function, z =fugacity; and, for a regular solution (assumed to have no vacant sites), (8)=g.p.f., z =ratio of the fugacities; and, for an Ising ferromagnet, (8)=partition function, $z=\exp(-2\mu H/kT)$ (μ =dipole moment, H =external field); excluding the unimportant additional factors.

the largest λ_i may become degenerate somewhere, where a phase transition (coexistence of states) occurs.¹¹⁾

Thus, in any case, if we want to discuss the phase transitions analytically (function-theoretically) in connection with mathematical singularities, we must treat *infinite* systems ($\alpha, \beta: V \rightarrow \infty$; $\gamma: l \rightarrow \infty$ and m or n or both $\rightarrow \infty$) and investigate the analytical behaviour of the rigorous limiting thermodynamic functions. We may consider *any* one of these limiting functions, since, if any one is singular, so also are all the others and a phase transition occurs. Thus, for example, we investigate the function φ (or $\bar{\varphi}$) of z (with T fixed).

§ 2.3. We shall make a classification of all possible types of analytical behaviours of the functions describing the condensing systems. We now propose the following questions about the function $\varphi(z)$ describing rigorously a condensing system of infinite size and about its singularity (P) representing the condensation.

Question 1: "Is the singularity *analytical*?" [Here by an *analytical* singularity we mean a singularity defined in the theory of functions of a complex variable, i.e., a singularity which an "analytic" function has, e.g., a pole, a branch-point, an essential singularity, etc. And any other kind of singularity should be called *non-analytical*, e.g., a point of intersection of two different analytic functions, and a point of intersection of two different branches of one analytic function;—see (b) and (d) of Fig. 2.1.]

Question 2: "Are the gaseous and the liquid phase described by one and the same analytic function?", i.e., "Is the function for the liquid phase given by the analytical continuation of the function for the gaseous phase?"

[Subsequently to the question 1, we propose *Question 1a*: "If the singularity is *analytical*, then is it *algebraic* or *transcendental*?", inquiring into the function-theoretical qualities of the singularity.*]

Answering both "Yes" and "No" to the questions 1 and 2 independently (see Table 2.1), we obtain the four types (a), (b), (c), and (d), all of which are *mathematically realizable* (see Fig. 2.1). [Metaphorically, if we mean by a *tree* one analytic function and by a *bamboo* another,** then (a) = the tree having a joint at P; (b) = the tree grafted on the tree at P; (c) = the bamboo grafted on the tree at a joint (P) of the tree; (d) = the bamboo grafted on the tree at P which is not a joint of the tree.] Consequently the questions 1 and 2 are independent of each other. [For example, define $\varphi(z)$ for $z > 0$, as the largest of the roots w of the irreducible algebraic equation

$$w(w+a-b)^2 - (z-c)^2 = 0 \quad (b > a > 0, c > 0) \quad (11)$$

(see Fig. 2.2); then this is a case of the type (b). The point $(c, b-a)$ is not a branch-point but a point of intersection of two different branches of the analytic (algebraic) function $w(z)$, and is therefore a non-analytical singularity of the function $\varphi(z)$ —GPL

* Only the type (a) necessitates *transcendancy*, by the discontinuity of the derivative at P. Any other type may be *algebraic* or *transcendental*.

** This metaphor is an extension of Katsura and Fujita's phraseology.¹³⁾

| | | | |
|-----|-----|-----|----|
| 1 | | Yes | No |
| 2 | | | |
| Yes | (a) | (b) | |
| No | (c) | (d) | |

Table 2.1

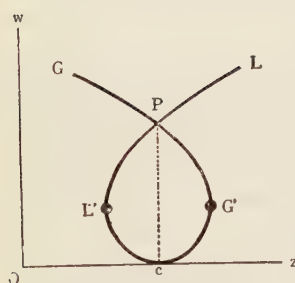
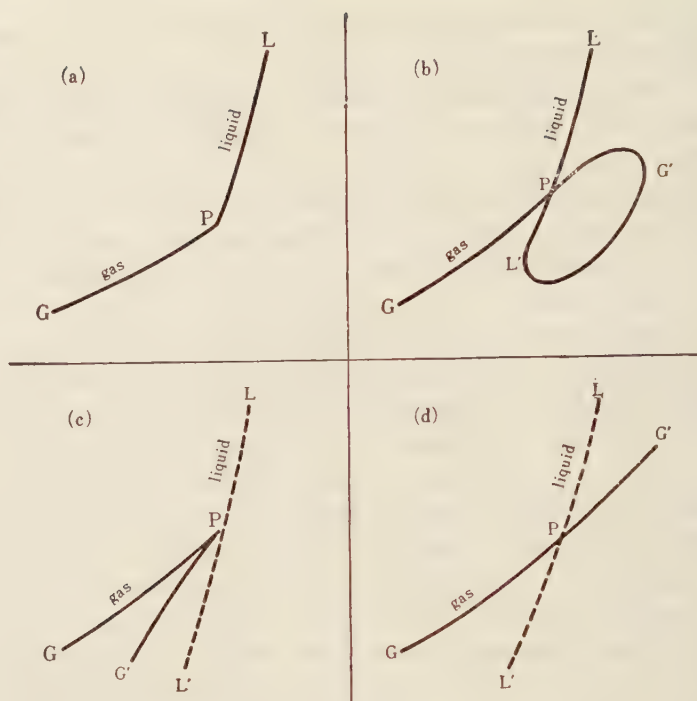


Fig. 2.2.

Fig. 2.1 Ordinate= φ . Coordinate= z [=real].

Function $\varphi(z) = GPL$ [=real], where GP=gaseous phase, PL=liquid phase, and P=condensation point.

A solid curve represents an analytic function, and a broken curve represents an analytic function different from that which a solid curve represents.

Some branches of the curves—e.g. the branch PG' in (c)—may be real or complex. The same is true in Fig. 2.3.

which is not an analytic function; thus it is the condensation point P. The points G' and L' are analytical singularities (algebraic branch-points) of $w(z)$.]

The classification into the four types (a), (b), (c), and (d) is essential because the question as to the analytical properties of the condensation point and the question as to the analytical relation between the functions for the gaseous phase and those for the liquid phase are both important from the analytical viewpoint; the latter question is also connected with the problem how to derive the rigorous functions for the liquid phase, since, if (a) or (b) is true, we can derive the liquid functions by prolonging (analytically) the gas functions.

The types (a), (b), (c), and (d), given in Fig. 2.1, are basic and standard; we may have many types by further classification and each type may be of various forms, as shown in Fig. 2.3.

The classification into the basic types (a), (b), (c), and (d) has been made by taking the gaseous phase as the starting-point and considering whether the singularity of the

function for the gaseous phase, that is, the point of "condensation", is an analytical or a non-analytical singularity. On the other hand, if we take the liquid phase as the

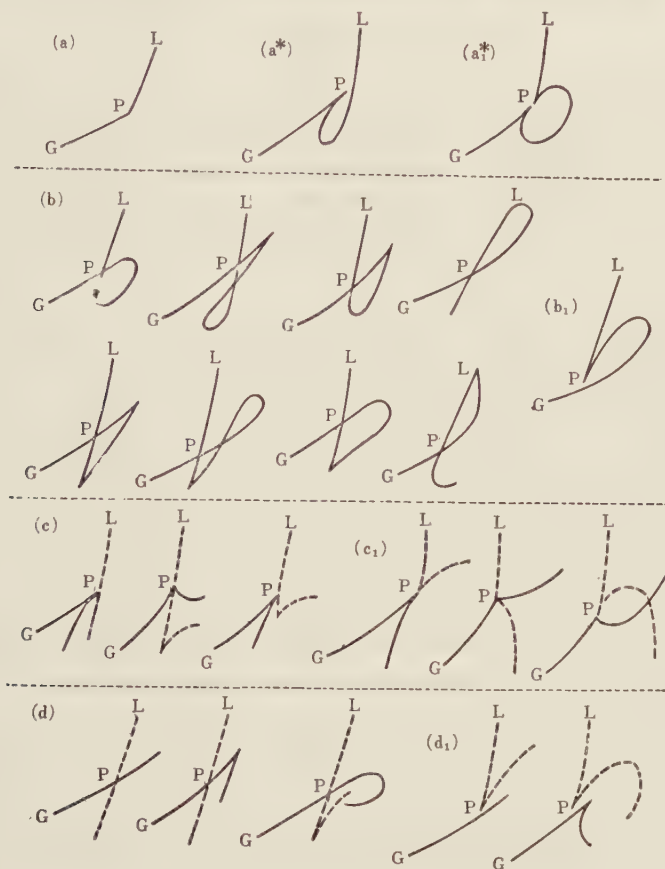


Fig. 2.3.

startingpoint, then "condensation" is replaced by "gasification" (although the point of condensation and the point of gasification are identical in position) and the basic types (a), (b), (c), and (d) are replaced by the basic types (a), (b), (d₁), and (d) as shown in Fig. 2.3. We usually take the gaseous phase as the starting-point, because the functions for the gaseous phase are comparatively easy to derive or discuss from the mathematical viewpoint.

§ 2.4. On the basis of the above stated conceptions and discussions, we shall now examine several authors' arguments on phase transitions.

§ 2.4.1. Mayer and others' theory⁽⁴⁾⁽⁵⁾⁽⁶⁾ by the method α and also Mayer's theory⁽⁸⁾ by the method β make answer "Yes" to the question 1, but give no answer to the question 2, since no functions for the liquid phase are derived or discussed. Thus we may say that Mayer and others⁽⁴⁾⁽⁵⁾⁽⁶⁾⁽⁸⁾ assert the truth of either the type (a) or (c) for real systems.

But here Mayer and others⁴⁾⁵⁾⁶⁾ use (for the gaseous phase and the transition range) the functions $\varrho_N^{(0)}(v, T)$, $\omega^{(0)}(v, T)$ and $\varphi_c^{(0)}(v, T)$ i.e.

$$\varphi_c^{(0)} \equiv \varphi_c^{(0)}(v, T) \equiv \lim_{N \rightarrow \infty} \varphi_{c,N}^{(0)}(Nv, T) \quad (12)$$

[where the superscript (0) means that every $b_i(V, T)$ in the functions without the superscript — Eqs. (1) and (2) — is replaced by $b_i^{(0)}(T)$, (ii) of Definition 1 (of Part I)] and show that the condensation occurs at the first singularity z (analytical) of $\varphi_c^{(0)}$ (and $\omega^{(0)}$) on the positive real axis* of z ; and Mayer⁴⁾ uses (for the gaseous phase) the function $\varphi_g^{(0)}(z, T)$ [where the superscript (0) has the same meaning as above — see Eqs. (5) and (6) —] and considers the first singularity z_c (analytical) of $\varphi_g^{(0)}$ on the positive real axis of z to be the condensation point.

The functions $\varphi_c^{(0)}$ and $\varphi_g^{(0)}$ should be distinguished from the functions φ_c and φ_g . Here $\varphi_c^{(0)}$ and $\varphi_g^{(0)}$ are shown to be identical (at least up to z_c), and will be denoted by $\varphi^{(0)}$. We have yet to discuss the question whether or not Mayer and others' assertion, based on $\varphi^{(0)}(\varphi_c^{(0)}, \varphi_g^{(0)})$, is correct for $\varphi(\varphi_c, \varphi_g)$ — for real systems.

§ 2.42. Katsura and Fujita¹³⁾ remark that φ_c and $\varphi_g^{(1)}$ in the method β — should be distinguished from each other, although they are equal for sufficiently small z , and conjecture that the first singularity (on the positive real axis of z) — the condensation point — of φ_g and that of $\varphi_g^{(0)}$ are different, and hence the former is a non-analytical singularity, which makes answer "No" to the question 1 and means the truth of either the type (b) or (d) for real systems, contrary to Mayer's assertion. We have yet to discuss the question whether or not Katsura and Fujita's conjecture is correct for real systems.

§ 2.43. S. Ono¹⁴⁾¹⁶⁾ attempts to discuss Katsura and Fujita's conjecture, from the direction of the method γ — crystal statistics. He has argued on the analytical behaviour of the roots of the secular equation (7). Now from our viewpoint, we shall analyze Ono's argument** and show that it cannot lead us to any conclusions about Katsura and Fujita's conjecture or Mayer's assertion.

When $m \times n$ is finite, the roots of (7) form one or more algebraic functions of z , and the function $\bar{\varphi}_{mn}(z)$ given by (9) is regular on the positive real axis of z , since the largest root of (7) is simple anywhere on the axis. On making $m \times n \rightarrow \infty$, the roots of (7) may, in general, form not only algebraic functions but transcendental functions,

* Here the variable $z = z(v, T)$ is defined by (7) of Lemma 6 (of Part I), and is proved to be the "fugacity" of the system.

** The opinion of Ono's first paper,¹⁴⁾ that if the secular equation (7) for $m \times n = \infty$ is irreducible (i.e. its roots form one analytic function) then the point of condensation is an analytical singularity (i.e. Mayer's assertion is correct), has been shown to be erroneous (and revised in ref. 16)) by the present author's remark¹⁵⁾ that even if (7) for $m \times n = \infty$ is irreducible, the realization of the type (b), as well as the type (a), is possible and so the point of condensation is not necessarily an analytical singularity, that is to say, the reducibility or irreducibility of (7) is of no importance to the analytical properties of the singularity representing the condensation. [See p. 85 of ref. 15) and p. 8 of ref. 16).]

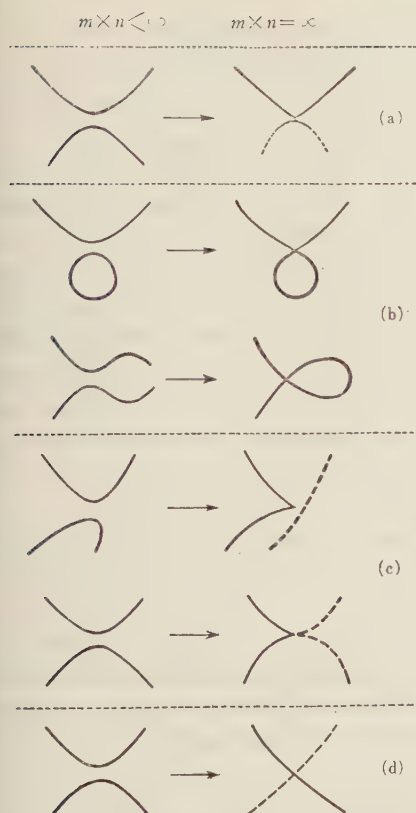


Fig. 2.4.

and the limit $\bar{\varphi}(z)$ of the function $\bar{\varphi}_{mn}(z)$ may have a singularity (analytical or non-analytical) on the real axis in such ways as are shown in Fig. 2.4 for the types (a), (b), (c), and (d).

That (7) for $m \times n = \infty$ should be irreducible, i.e., that the roots of (7) for $m \times n = \infty$ should form one analytic function, is a sufficient (but not necessary) condition that one should make answer "Yes" to the question 2, i.e., the type (a) or (b) should be true.

On the other hand if (7) for $m \times n = \infty$ is reducible, i.e., the roots of (7) for $m \times n = \infty$ form two or more analytic functions, i.e., the infinite determinant is divided into a finite or infinite number of irreducible factors, then each of these factors is a finite determinant or else an infinite determinant (the former gives an algebraic function as a part of the set of roots of (7) while the latter may give a transcendental function). Then, if the function $\bar{\varphi}(z)$ ($z > 0$) is given by only one of the factors (a finite or infinite determinant), one makes answer "Yes" to the question 2, i.e., the type (a) or (b) is true. But if the function $\bar{\varphi}(z)$ ($z > 0$) is given by two (or more)

of the factors, one makes answer "No" to the question 2, i.e., the type (c) or (d) is true.

Thus further discussions on the reducibility or irreducibility* of (7) might, in some way or other, be related to the question 2. But we have no way of answering the question 1 from this direction.

§ 2.44. Yang and Lee's theory (part I)¹⁸⁾ of phase transitions utilizes the method β and seems to argue that answer "Yes" is given to the question 1; but this argument will be shown to be incorrect, according to the discussion in § 3.4 of the present paper.

§ 2.5. Thus we assert that any discussion on phase transitions from the analytical viewpoint should first give an answer to the questions 1 and 2, that is to say, it should determine which type of (a), (b), (c) and (d), is realized. In the following we shall go on along these lines, applying our lemmas and theorems.

* In fact, even if (7) is irreducible for $m \times n < \infty$, it is somewhat possible that it may be reducible for $m \times n = \infty$, since some elements of the matrix vanish in this limit.

Part III

Theory of Condensation

§ 3.0. In this article, using some of the results of Parts I and II, we shall attempt to discuss the condensation phenomena of real systems. Here we shall see an example in which a theorem of Part I is applied to a physical problem.

§ 3.1. The irregular points of the limiting function ω [Eq. (9) of § 0.1] represent the points of phase transition of the "real system", while the irregular points of the limiting function $\omega^{(0)}$ [Eq. (7) of § 0.1] represent the points of phase transition of the "(0)-system"; in particular, the irregular point of the function ω^\dagger of v (or z), furthest from the origin of v (or nearest to the origin of z), represents the point of condensation of the "real system", while the irregular point of the function $\omega^{(0)}$ of v (or z), furthest from the origin of v (or nearest to the origin of z), represents the point of condensation of the "(0)-system".*

In Part I, we have discussed under what conditions the equality [Eq. (10) of § 0.2]

$$\omega^\dagger = \omega^{(0)} \quad (1)$$

is true. Now we regard T as a parameter, which is fixed at a value; then ω^\dagger and $\omega^{(0)}$ are considered as functions of v . Then, since we have $z = z(v, T)$ by the definition of z [Eq. (7) of Lemma 6 of § 1.4], ω^\dagger and $\omega^{(0)}$ are considered as functions of z . Thus $\omega^\dagger = \omega^\dagger(v, T)$ or $\omega^\dagger = \omega^\dagger(z, T)$, and $\omega^{(0)} = \omega^{(0)}(v, T)$ or $\omega^{(0)} = \omega^{(0)}(z, T)$; $v > 0$, $z > 0$,

* By (3) of § 0.1, ω^\dagger and $\omega^{(0)}$ are the configurational parts of the Helmholtz free energies per molecule for infinite systems.

Strictly speaking, we have yet to prove

$$\lim_{N \rightarrow \infty} \varphi_{e, N}(Nv, T) = (\partial/\partial v) \omega(v, T) \quad (i)$$

where [cf. (1) of § 2.1]

$$(p/kT) \varphi_{e, N}(V, T) = (\partial/\partial V) \ln \mathcal{Q}_N(V, T) \quad (ia)$$

$$\omega(v, T) = \lim_{N \rightarrow \infty} (1/N) \ln \mathcal{Q}_N(Nv, T). \quad (ib)$$

This may be easily proved if one extends the lemmas and theorems of Part I—stated for any fixed value of v (or z)—to the case of a region of values of v (or z) and employs the well-known "Weierstrass theorem on double-series" in a modified form, since it may be proved that the infinite sequences which appear in our lemmas and theorems and in their proofs are uniformly convergent in any region interior to the region (of v or z) in question.

Then, if $\omega^\dagger = \omega^{(0)}$ for a region, we have for this region

$$p_{\omega^\dagger} = p_{\omega^{(0)}} \quad (ii)$$

where

$$p_{\omega^\dagger} = kT(\partial/\partial v) \omega^\dagger \quad (iia)$$

$$p_{\omega^{(0)}} = kT(\partial/\partial v) \omega^{(0)} \quad (iib)$$

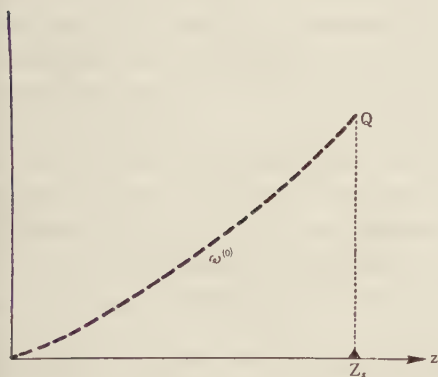
where, by (i), p_{ω^\dagger} and $p_{\omega^{(0)}}$ are the pressures for infinite systems. If and only if ω is irregular, p_{ω^\dagger} is irregular. The same is true for any other thermodynamic function.

and $T > 0$. From the definition of z it is proved that z is the *fugacity* (or *activity*) of the “(0)-system”. If the equality (1) is true, then the fugacity of the “real system” is equal to the fugacity z of the “(0)-system”, that is to say, z means the fugacity of both the “(0)-system” and the “real system”.

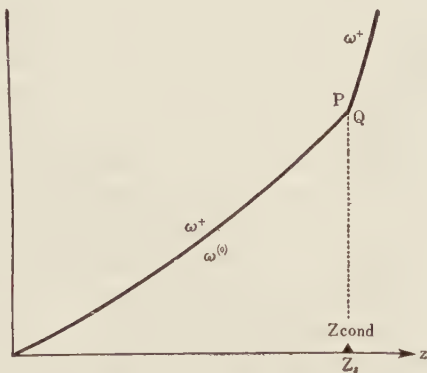
§ 3.11. From Argument 1 of § 1.4 it follows that the function $\omega^{(0)}$ of z is regular for all z less than z_s and has a singularity at z_s [see Fig. 3.1 (i)], and that this singularity is an “analytical” singularity [where the word “analytical” is defined in § 2.3]. Thus the condensation for the “(0)-system” occurs at z_s , an “analytical” singularity.

§ 3.12. If the equality (1) is true for all z less than z_s [see Fig. 3.1 (ii)], then the function ω^\dagger of z is regular for all z less than z_s and has a singularity at z_s (since $\omega^{(0)}$ and ω^\dagger are the same analytic function for z less than z_s), and so the condensation

(i)



(ii)



(iii)

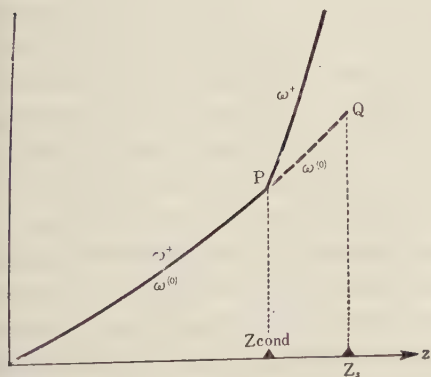


Fig. 3.1 [Analytical behaviour of the thermodynamic functions of condensing systems] (i): the (0)-system. (ii): the real system (case 1). (iii): the real system (case 2). P represents the singularity of ω^\dagger that is the condensation point of the real system. Q represents the singularity of $\omega^{(0)}$ that is the condensation point of the (0)-system.

for the “real system” occurs at z_s ; that is, z_{cond} (i.e., the condensation point of the real system) $= z_s$; and this singularity z_s of the function ω^\dagger is “analytical”. Hence, in

this case, the condensation for the "real system" occurs at an "analytical" singularity, and Mayer's theory about the connection between the singularity and the condensation is correct even if we rigorously consider the volume dependence of the cluster integrals. That is, we make answer "Yes" to the question 1 of § 2.3:—the type (a) or (c) is true.

§ 3.13. On the contrary if the equality (1) is true for all z less than a value of z (z_{cond} say) less than z but is not true beyond the point z_{cond} [see Fig. 3.1 (iii)], then the function ω^1 of z is regular for all z less than z_{cond} and has a singularity at z_{cond} , and so the condensation for the "real system" occurs at z_{cond} ($\cdot z$), and this singularity z_{cond} of the function ω^1 is "non-analytical". Hence, in this case, we make answer "No" to the question 1 of § 2.3:—the type (b) or (d) is true; and Mayer's theory does not hold.

§ 3.14. Similar statements are true with v , v_s , v_{cond} in place of z , z_s , z_{cond} and with "greater" in place of "less" in §§ 3.11, 3.12 and 3.13, since [by Argument 1 of § 1.4] the function $z(v, T)$ of v is regular (and decreasing) for $v > v_s$ (i.e., for $z < z_s$) and has an "analytical" singularity at v_s . Here $z_s = z(v_s)$ and $z_{\text{cond}} = z(v_{\text{cond}})$.

§ 3.2. Conditions for the truth of the equality (1) are given in Part I by Theorem α and its Corollaries 1 and 2 (of § 1.5).

The Corollary 2 states a sufficient condition for the truth of the equality (1). It should be remarked that the condition for the equality has been expressed by inequalities. It is much easier to examine the truth of an inequality than that of an equality. In this meaning the corollary, as well as the theorem, may be very useful.

Using the corollary, we attempt a discussion, as follows, of the problem as to condensation and singularity.*

To confirm that the inequalities are satisfied, we need only make a rough estimation of the behaviour of the "real" cluster integrals $b_l(V, T)$ as functions of the volume V . This purpose may be effected, for example, by examining the behaviour of $b_l(V, T)$ for a lattice gas of small size, that is, a system consisting of L lattice points and N particles [where L and N ($\leq L$) are small and $V = LC$ (C being a suitable constant)], since such a system may be considered to describe roughly the behaviour of $b_l(V, T)$ of a real large system. Such an examination has been given, for example, by the work¹⁷⁾ of Fujita and Katsura. For all sufficiently low temperatures it is shown that the function $b_N^1(Nv, T)$ of v is positive and increasing with v for all values of v greater than a value (v_0 say) of v . This is due to the predominance of the positive part (attractive part) of the function $f(r_{ij})$ in the integrand of the cluster integral for low temperatures and to the increase of the integration region of the cluster integral with increasing v . For $v < v_0$, $b_N^1(Nv, T)$ may be negative or positive and its behaviour is, in general, very irregular. This is due to the irregular contribution of the negative part (repulsive part) of the function $f(r_{ij})$ in the integrand of the cluster integral for small volumes. It is seen

* It may be noted that the discussion in this section (§ 3.2) is not so rigorous as the proofs in Part I and so there is some difference in character between the argument in Part I and that in this section.

that v_0 is small as compared with v_s (i.e., the specific volume for the saturated vapour), and that $b_l^{(0)}(T) > b_l^l(Nv, T) > b_{lN}^l(Nv, T)$ for $l \leq N$ and for $v \geq v_0$. Thus we may consider that the condition (C) of Corollary 2 (of § 1.5) is satisfied for all $z < z_s$ (i.e., for all $v > v_s$) at all temperatures so low that the usual condensation occurs, and hence, at all such temperatures, Mayer's theory of condensation is true and the condensation point of the "real system" is an "analytical" singularity.

In the present theory we are restricted to the range of temperatures for which every $b_l^{(0)}(T)$ is positive, that is, the range of temperatures lower than some temperature which may be a little lower than the critical temperature (see [[3]] of Argument 1 of § 1.4). [Condensation phenomena of usual character, as treated by Mayer's original theory¹⁾ (about the condensation of the "(0)-system") in terms of usual "liquid drop clusters", are considered to occur in this range of temperatures.] But some $b_l^{(0)}(T)$ may be negative for not so low temperatures, that is, for temperatures sufficiently near the critical temperature. Moreover, for temperatures for which every $b_l^{(0)}(T)$ is positive but which is sufficiently near the critical temperature, the condition of Corollary 2 (of § 1.5) and even the condition of Theorem [I] (of § 1.5) may be invalid for some values of v such that $v > v_s$. Consequently, in any case, near the critical temperature, the condensation point of the "real system" may be a "non-analytical" singularity. Thus, for this range of temperatures, the condensation phenomenon of the "real system" may have a sort of unusual or irregular character. Perhaps this might be connected with the observed singular phenomena of a gas in the critical region, i.e., the region near the critical point.

§ 3.3. In Part I we have obtained also Theorem β (of § 1.6), which states conditions for inequality and may be useful in the theory of condensation.

In case $v_{cond} > v_s$ (i.e., in case Mayer's theory be invalid), either the condition of Theorem [III] (of § 1.6) or the condition of Theorem [IV] (of § 1.6) should be satisfied for all v such that $v_{cond} > v > v_s$, to ensure the existence of the limiting function from which the thermodynamic functions (per molecule) of the "real system" should be derived.

§ 3.4. Here we shall compare the present theory with Yang and Lee's theory¹⁸⁾ of phase transition. While the present theory concerns the case of canonical ensemble, Yang and Lee consider the case of grand canonical ensemble and attempt to prove the truth of the equality [Eq. (16) of reference 18)]

$$\lim_{t \rightarrow \infty} \sum_{l=1}^{\infty} b_l^l(V, T) z^l = \sum_{l=1}^{\infty} b_l^{(0)}(T) z^l. \quad (2)$$

The left-hand side of Eq. (2) describes the "real system", and so it corresponds to the function ω^l for our case. The right-hand side of Eq. (2) describes the "(0)-system", and so it corresponds to the function $\omega^{(0)}$ for our case. [Of course, each side of Eq. (1) (for our case) is related directly to the Helmholtz free energy per molecule, but each side of Eq. (2) (for Yang-Lee's case) to the pressure. However, in the theory of condensation, one may discuss the analytical behaviour of any thermodynamic function.] Hence the equality (2) for Yang-Lee's case corresponds to the equality (1) for our case. Yang and Lee have proved the truth of the equality (2) for all z less than t_1 ; [this t_1

is seen to be z_{cond} for our case]. But in Yang-Lee's argument the following two different concepts have been confused by using the same symbol t_1 for them.

(I): The first singularity along the positive real axis, of the series $\sum_{l=1}^{\infty} b_l^{(0)}(T)z^l$ or its analytical continuation.

(II): The point at which the roots of $\Xi_V(z)=0$ close in onto the real axis as $V \rightarrow \infty$, (where $\Xi_V(z)$ is the grand partition function).

Both (I) and (II) have been written as t_1 by Yang and Lee. However, (I), which means the condensation point of the "(0)-system", should be written as $t_1^{(0)}$ (i.e. z_s), and (II), which means the condensation point of the "real system", should be written as t_1^{\dagger} (i.e. z_{cond}). In Yang-Lee's argument, $t_1^{(0)}$ and t_1^{\dagger} have not been proved to be equal.

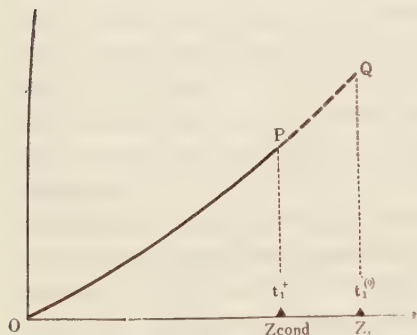


Fig. 3.2 [Comparison between the arguments of Yang-Lee and of ours] The ordinate represents both the Helmholtz free energy (per molecule) and the pressure, for the real system or for the (0)-system.

Now see Fig. 3.2. We may say that Yang and Lee have proved the truth of the equality (2) for the part OP, that is, up to z_{cond} (i.e. t_1^{\dagger}); this does not mean the proof of the truth of Mayer's theory of condensation, because it is possible that P, the condensation point of the "real system", and Q, that of the "(0)-system", may be different, that is, $z_{\text{cond}} < z_s$ (i.e. $t_1^{\dagger} < t_1^{(0)}$). On the other hand, in the present theory (§ 3.2), (for all temperatures such that the usual condensation occurs) we have attempted

to confirm the truth of the equality (1) for the part OQ (i.e., up to z_s), which means the coincidence* between P and Q (i.e., $z_{\text{cond}} = z_s$) and consequently the truth of Mayer's theory of condensation.

Acknowledgement

The author wishes to express his sincere thanks to Prof. K. Husimi (Osaka University), Prof. J. E. Mayer (Chicago University), Dr. T. Tanaka (Kyusyu University) and many others for comments and discussions made when the present paper was read at the International Conference of Theoretical Physics (1953).

Appendices (to Part I)

Appendix 1. «Proof of Lemma γ »

(i) and (ii) of Lemma γ are obvious. Let us prove (iii).

1° If $\max_{l \geq 1} \theta(l, T)$ exists, there is the least value (say l_1) of l' such that $\theta(l', T) = \max_{l \geq 1} \theta(l, T)$. Then $\bar{\theta}(l, T) = \max_{l' \geq 1} \theta(l', T)$ for all $l \geq l_1$. Hence (a'). Now $\theta(l, T)/\bar{\theta}(l, T) \leq 1$ for all $l \geq l_0$, and

* Note that, since z_s (i.e. $t_1^{(0)}$) is an "analytical" singularity, it is of course impossible that $z_s < z_{\text{cond}}$ (i.e. $t_1^{(0)} < t_1^{\dagger}$).

$\theta(l, T)/\bar{\theta}(l, T)=1$ and $l_1 \geq l_0$. Hence (b').

2° If $\max_{l \geq 1} \theta(l, T)$ does not exist, then $\bar{\theta}(l, T) = \theta(l, T)$ for an infinite number of l . For, if there be only a finite number of l such that $\bar{\theta}(l, T) = \theta(l, T)$, then there is the greatest (say l_2) of such l ; and $\theta(l, T) < \theta(l_2, T)$ for all $l > l_2$, and $\theta(l, T) \leq \theta(l_2, T)$ for all $l \leq l_2$; hence $\theta(l_2, T)$ is greatest, contrary to the assumption. Therefore $\lim_{l \rightarrow \infty} \bar{\theta}(l, T)$ is a limit-value of the sequence $\{\theta(l, T)\} (l=1 \dots \infty)$. But $\bar{\theta}(l, T) \geq \theta(l, T)$ for all l , and so the sequence $\{\theta(l, T)\} (l=1 \dots \infty)$ has no limit-values greater than $\lim_{l \rightarrow \infty} \bar{\theta}(l, T)$. Hence (a''). Since $\theta(l, T)/\bar{\theta}(l, T)=1$ for an infinite number of l , unity is a limit-value of $\{\theta(l, T)/\bar{\theta}(l, T)\} (l=l_0 \dots \infty)$. But $\theta(l, T)/\bar{\theta}(l, T) \leq 1$ for all $l \geq l_0$, and so $\{\theta(l, T)/\bar{\theta}(l, T)\} (l=l_0 \dots \infty)$ has no limit-values greater than unity. Hence (b'').

Appendix 2. « Lemma »

For any positive integer l' and for every integral value of s such that $0 \leq s \leq (1/2)l'(l'+1)$, we can find a set $\{\nu_l\} (l=1, 2, \dots, l')$ such that:

1° each ν_l is either 0 or 1,

2° $\sum_{l=1}^{l'} l\nu_l = s$.

(Proof) — Take any integer s' such that $0 \leq s' \leq (1/2)l'(l'+1) - 1$, and assume that we can find a set $\{\nu_l\}$ which satisfies 1° and 2° for $s=s'$. In this set there is at least one l such that $\nu_l=0$. [For, if $\nu_l=1$ for every $l \leq l'$, then $s'=(1/2)l'(l'+1)$, contrary to the assumption.] Let m be the least of such l . If $m=1$, change ν_1 from 0 to 1; if $m \geq 2$, change ν_m from 0 to 1 and ν_{m-1} from 1 to 0. Then in either case we get a set $\{\nu_l\}$ which satisfies 1° and 2° for $s=s'+1$. The set $\{0, 0, \dots, 0\}$ satisfies 1° and 2° for $s=0$. Hence the lemma.

Appendix 3. « Notes for the proof of Lemma 6 »

[Here the indices 1, 2, 3, ..., 9, A, B, C correspond to the superscripts of (28)³, (31b)^C, etc., in Lemma 6, and (n) is the number of a formula in Lemma 6.]

From (11a) $\hat{m}_l = [\bar{m}_l] + \mu_l$, (17) $\bar{m}_l > 0$, and (20) $\mu_l \geq 0$, we have:

1. $0 \leq \bar{m}_l - \hat{m}_l < 1$ (if $\mu_l = 0$). 2. $0 < \hat{m}_l - \bar{m}_l \leq \mu_l$ (if $\mu_l \geq 1$). 3. $|\hat{m}_l - \bar{m}_l| < \mu_l + 1$, (from 1 and 2). 4. $\mu_l \geq 1$ (if $\hat{m}_l \geq 1$ and $\bar{m}_l < 1$). 5. $\bar{m}_l \geq 1$ (if $\hat{m}_l \geq 1$ and $\mu_l = 0$). 6. $0 < \bar{m}_l < 1$ (if $\hat{m}_l = 0$).

From (13), (15) and (19), we have 7. $0 \leq \mu_l \leq \mu(N) + 1$ for all $l \leq l'(N)$ and for all N .

From (14), (17) and (19), we have 8. $0 < \mu(N) \leq N$ for all N .

From (21) and (22), we have 9. $0 \leq \hat{m}_l \leq N$ for all $l \leq l'(N)$ and for all N .

Now we have the following lemmas A, B, and C.

A. If x is real and $x \geq 1$ and μ is a non-negative integer, then we have

$$|\ln\{([x] + \mu)/x\}| < \ln(\mu + 2).$$

[This follows from the following two lemmas A_1, A_2 .]

A_1 . If x is real and $x \geq 1$, then $0 \leq \ln([x]/x) > -\ln 2$.

A_2 . If x, μ are real and $x \geq 1, \mu \geq 1$, then

$$0 < \ln\{([x] + \mu)/x\} \leq \ln(\mu + 1).$$

B. The function $x \ln([x]/x)$ is bounded in the half-open real interval $1 \leq x < +\infty$.

[For we have $x \ln(1-1/x) < x \ln([x]/x) \leq 0$ for $x > 1$; and $x \ln(1-1/x)$ is bounded for $x \geq 2$, since it is continuous for $x \geq 2$ and its limit for $x \rightarrow +\infty$ is -1 . And $x \ln(1/x)$ is bounded for $1 \leq x < 2$.]

C. If x, μ are real and $x > 0, \mu > 0$, then we have

$$0 < x \ln\{([x] + \mu)/x\} < \mu.$$

[For we have $0 < x \ln(1 + \mu/x) < \mu$ ($x > 0, \mu > 0$), since $\lim_{x \rightarrow 0} x \ln(1 + \mu/x) = 0$ and $\lim_{x \rightarrow +\infty} x \ln(1 + \mu/x) = \mu$ and

$$(d/dx)\{x \ln(1 + \mu/x)\} = 1/(1 + \mu/x) - \ln\{1/(1 + \mu/x)\} - 1 > 0 \quad (x > 0, \mu > 0).]$$

Appendix 4. «Auxiliary theorems»

I. For any infinite sequence $\{a_n\}$ ($n=1, \dots, \infty$) of real numbers, the following are true.

- (i) If and only if $\limsup a_n = \lambda$ ($-\infty < \lambda < +\infty$), then, given any $\epsilon > 0$, we have $a_n < \bar{\lambda} + \epsilon$ for almost all n , and we have $a_n > \bar{\lambda} - \epsilon$ for an infinite number of n .
- (ii) If and only if $\limsup a_n = +\infty$, then, given any real K (however large it may be), we have $a_n > K$ for an infinite number of n .
- (iii) If and only if $\limsup a_n = -\infty$, then, given any real k (however small it may be), we have $a_n < k$ for almost all n .

II. For any two sequences $\{a_n\}$ and $\{b_n\}$, if $a_n \leq b_n$ for almost all n , then $\limsup a_n \leq \limsup b_n$.

III. For any two sequences $\{a_n\}$ and $\{b_n\}$ and for $\{c_n = a_n + b_n\}$, we have $\liminf a_n + \liminf b_n \leq \liminf c_n \leq \limsup a_n + \limsup b_n$.

[Note] The cases that $(+\infty) + (-\infty)$ and that $(-\infty) + (+\infty)$, and only these cases, should be excluded. The same is true in the following corollaries.

Corollary 1. If $\lim a_n$ exists, then, (i) $\limsup c_n = \lim a_n + \limsup b_n$, (ii) $\liminf c_n = \lim a_n + \liminf b_n$. [See III and note that $\limsup a_n = \liminf a_n = \lim a_n$.]

Corollary 2. If $\{a_n\}$ is convergent, then $\{b_n\}$ and $\{c_n\}$ are convergent or divergent (to positive or negative infinity) or oscillatory, together. [By Corol. 1; if $-\infty < \limsup b_n = \liminf b_n < +\infty$, then $-\infty < \limsup c_n = \liminf c_n < +\infty$; if $\limsup b_n = \liminf b_n = \pm\infty$, then $\limsup c_n = \liminf c_n = \pm\infty$; if $\limsup b_n \neq \liminf b_n$, then $\limsup c_n \neq \liminf c_n$.]

Corollary 3. If both $\lim a_n$ and $\lim b_n$ exist, then $\lim c_n = \lim a_n + \lim b_n$. [See Corol. 1 and note that $\limsup b_n = \liminf b_n = \lim b_n$.]

References

- 1) K. Ikeda, *Busseiron Kenkyu* No. 65 (1953), 145.
- 2) K. Ikeda, *Proc. Internat. Conf. Theor. Phys. Kyoto & Tokyo* (1953), 544.
- 3) H. D. Ursell, *Proc. Camb. Phil. Soc.* **23** (1927), 685.
- 4) J. E. Mayer, *J. Chem. Phys.* **5** (1937), 67.
J. E. Mayer & Ph. G. Ackermann, *ibid.* **5** (1937), 74.
J. E. Mayer & S. F. Harrison, *ibid.* **6** (1938), 87.
S. F. Harrison & J. E. Mayer, *ibid.* **6** (1938), 101.
J. E. Mayer & M. G. Mayer, *Statistical Mechanics* (J. Willey & Sons, 1940), Chapters 13 and 14.
- 5) M. Born & K. Fuchs, *Proc. Roy. Soc. A* **166** (1938), 391.
- 6) B. Kahn & G. E. Uhlenbeck, *Physica* **4** (1938), 299.
- 7) K. Ikeda, *Busseiron Kenkyu* No. 52 (1952), 21.
K. Ikeda, *Prog. Theor. Phys.* **11** (1954), 336.
- 8) J. E. Mayer, *J. Chem. Phys.* **10** (1942), 629.
- 9) H. A. Kramers & G. H. Wannier, *Phys. Rev.* **60** (1941), 252, 263.
- 10) E. N. Lassetre & J. P. Howe, *J. Chem. Phys.* **9** (1941), 747, 801.
- 11) E. Montroll, *J. Chem. Phys.* **9** (1941), 706; **10** (1942), 61.
- 12) Y. Muto, *J. Phys.-Math. Soc. Japan* **17** (1943), 86.
Y. Muto, *J. Chem. Phys.* **16** (1948), 519, 524.
- 13) S. Katsura & H. Fujita, *Prog. Theor. Phys.* **6** (1951), 498.
S. Katsura & H. Fujita, *J. Chem. Phys.* **19** (1951), 795.
- 14) S. Ono, *Busseiron Kenkyu* No. 38 (1951), 1.
- 15) K. Ikeda, *Busseiron Kenkyu* No. 57 (1952), 77.
- 16) S. Ono, *Prog. Theor. Phys.* **8** (1952), 1 (revised after Ikeda's discussions¹⁵⁾).
- 17) H. Fujita & S. Katsura, *Sci. Rep. Tohoku Univ. Ser. I*, **35** (1952), 206.
- 18) C. N. Yang & T. D. Lee, *Phys. Rev.* **87** (1952), 404 (Part I).

Erratum

[K. Ikeda, *Proc. Internat. Conf. Theor. Phys. Kyoto & Tokyo* (1953), 544.]
p. 544, line 11. For $f_{ij} = \exp\{-u(r_{ij})/kT\}$ read $f_{ij} = \exp\{-u(r_{ij})/kT\} - 1$.

Nuclear Reactions in the Later Stage of the Stellar Evolution

Kimiko NAKAGAWA, Takashi OHMURA,[†] Hisao TAKEBE,
and Shinya OBI*

Department of Physics, University of Tokyo, Tokyo

**College of General Education and Astronomical Observatory, University of Tokyo, Tokyo*

(Received May 28, 1956)

Nuclear reactions induced by helium capture, $3\alpha \rightarrow \text{C}^{12} + \gamma$ etc., are investigated. The results are applied to the study of the variation of chemical abundances and the energy generation in the hydrogen exhausted stellar core with suitable physical conditions. The reactions, $\text{C}^{12} + \text{C}^{12}$ etc., are also discussed. The results obtained are useful to study the stellar evolution in its later stage.

§ 1. Introduction

Taketani, Hatanaka, and one of the present authors (Obi)¹⁾ have suggested that the characteristic differences between two populations of celestial objects are accounted for by the evolutionary scheme involving the building up process of heavier elements in the central region of stars of Population II and of Population I in the later stage of their evolution. The present paper aims to give detailed theoretical investigations of such synthesis processes of heavier elements in our evolutionary framework.**)

After the complete exhaustion of hydrogen in the central region of stars in the later stage of their evolution, the temperature and the density may rise high enough for the conversion of its helium into heavier elements. The nuclear reactions at this stage of stellar evolution were discussed by several authors²⁾³⁾⁴⁾.

In this paper the reaction rates of several helium reactons, $3\alpha \rightarrow \text{C}^{12} + \gamma$, $\text{C}^{12} + \alpha \rightarrow \text{O}^{16} + \gamma$, $\text{O}^{16} + \alpha \rightarrow \text{Ne}^{20} + \gamma$, are discussed in full detail. Our results differ from the earlier results obtained by Öpik and Salpeter mainly because of the adoption of new data on nuclear energy levels.

The detailed accounts are also given concerning the simultaneous occurring of these three helium reactions in the hydrogen exhausted core with particular reference to the stellar energy production. According to our numerical results, the energy production through these reactions is very large and is found to be sufficient to supply the whole radiation energy of very luminous stars such as those found in the globular clusters.

[†]) The name Kikuta has been changed to Ohmura.

***) A study on the same problem has been done also by S. Hayakawa, C. Hayashi, M. Imoto and K. Kikuchi the results of which will soon appear in this journal.

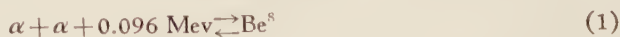
The reactions $\text{Ne}^{20} + \alpha$, etc. and $\text{C}^{12} + \text{C}^{12}$, etc. are also discussed, and are shown not to occur at any considerable rates under usual stellar conditions. It may be quite certain that under the conditions where the helium reactions are occurring above reactions have no appreciable effects on the energy production in usual stars. On the other hand, under the unusual conditions in the stellar collapsing stage or after the complete exhaustion of helium, these reactions may play a very important role especially in connection with the building up processes of heavy elements.

In the following, the rates of successive reactions induced by helium capture are discussed in § 2. The nuclear energy produced by the reaction $3\alpha \rightarrow \text{C}^{12}$ is calculated for several sets of the central temperature T_c and central density ρ_c of the helium convective core. This is shown in graphical form in § 3. In § 4 we formulate the simultaneous differential equations governing the procedures of these nuclear reactions in a stellar helium convective core. In § 5 using the typical giant stellar model, the differential equations obtained in § 4 are numerically solved and the change of chemical composition and stellar energy production with time are calculated. The reactions $\text{C}^{12} + \text{C}^{12}$ and some other reactions are discussed in § 6. In § 7, the obtained results are discussed and further remarks are given.

§ 2. The rates of reactions induced by helium capture

In a stage of the evolution of stars, the hydrogen in the active core will be gradually consumed through the reaction $p \rightarrow \alpha$. Öpik²⁾, Salpeter³⁾ and Hoyle⁴⁾ have shown that helium reactions will take place as the next stage. Öpik and Salpeter have paid special attention to the formation of C^{12} , and Hoyle has further calculated the rate of the formation of O^{16} in some detail. Recently, however, new experimental data on the energy levels of these nuclei have been presented. Therefore we have reinvestigated these reactions in detail, and have discussed also the formation of Ne^{20} which has been discussed by these authors only qualitatively. The rates of the further α -captures are small as shown in the latter half of this section.

When the temperature T and the density ρ of the central region of helium core become high enough, C^{12} is formed by the following processes:

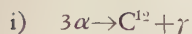


Once C^{12} is formed, the heavier elements will be produced by the successive capture of α -particle:



.....

To begin with, we shall calculate the rates of the above reactions (I), (II), (III), etc..



The scheme of this reaction is shown in Fig. (1)⁵⁾. In the following we shall discuss

in detail the reaction rate of this process.

Let each gram of the stellar material contain X_A gram of element A, and assume that the momentum distribution of α -particles is given by the Maxwell distribution law, namely the number of α -particles with momenta between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$ per unit volume is

$$n_\alpha (2\pi M_\alpha kT)^{-3/2} \exp \left[-p^2 / (2M_\alpha kT) \right] d^3\mathbf{p}, \quad (4)$$

where

$n_\alpha = \rho X_\alpha / M_\alpha$: number of α -particles per unit volume,

M_α : mass of α -particle,

k : Boltzmann's constant.

The cross sections of the reactions

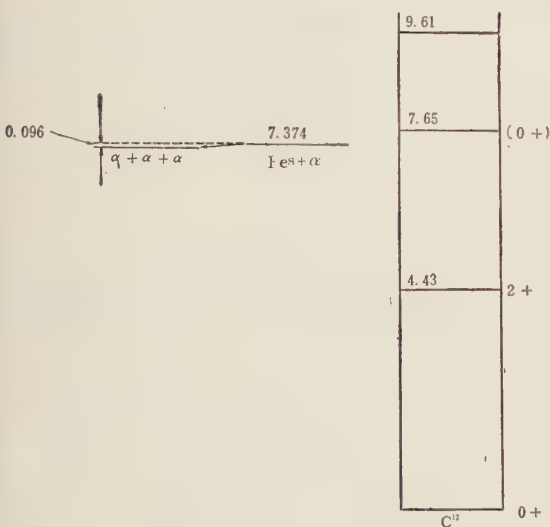


Fig. (1) Energy diagram of the reaction $3\alpha \rightarrow \text{C}^{12} + \gamma$ in MeV

$\alpha + \alpha \rightarrow \text{Be}^8$ and $\text{Be}^8 + \alpha \rightarrow \text{C}^{12} + \gamma$ are given by

$$\sigma^{R\alpha}(E) = (\pi \hbar^2 / 2M_I E) \left[\Gamma_\alpha^{R\alpha} \Gamma^{R\alpha} / \{ (E - E^{R\alpha})^2 + (\Gamma^{R\alpha}/2)^2 \} \right], \quad (5)$$

$$\sigma^c(E) = (\pi \hbar^2 / 2M_I' E') \left[\Gamma_\alpha^c \Gamma_\gamma^c / \{ (E - E^c)^2 + (\Gamma^c/2)^2 \} \right], \quad (6)$$

where

$$M_I = M_\alpha M_\alpha / (M_\alpha + M_\alpha) = M_\alpha / 2, \quad M_I' = M_\alpha M_{Be} / (M_\alpha + M_{Be}) = 2M_\alpha / 3,$$

$E^{R\alpha}$, E^c : resonance energies of the above reactions,

$\Gamma^{R\alpha}$, Γ^c : total widths of the ground state of Be^8 and the second excited state of C^{12} ,

$\Gamma_\alpha^{R\alpha}$, Γ_α^c : partial widths for α -emission of the ground state of Be^8 and the second excited state of C^{12} ,

Γ_γ^c : partial width for γ -emission of the second excited state of C^{12} .

As the life time of Be^8 is given by $\hbar / \Gamma^{R\alpha}$, from eqs. (4) and (5), the number of Be^8 per unit volume, n_{Be} is

$$n_{Be} = n_\alpha^2 (2\pi M_\alpha kT)^{-3} \int d^3\mathbf{p} \int d^3\mathbf{p}' (\hbar / \Gamma^{R\alpha}) (2E / M_I)^{1/2} (\pi \hbar^2 / 2M_I E) \\ \times \left[\Gamma_\alpha^{R\alpha} \Gamma^{R\alpha} / \{ (E - E^{R\alpha})^2 + (\Gamma^{R\alpha}/2)^2 \} \right] \exp \left[- (p^2 + p'^2) / (2M_\alpha kT) \right].$$

In this equation, we convert the variables from \mathbf{p} and \mathbf{p}' to the momentum \mathbf{P} of the centre of gravity and the momentum \mathbf{p}_r of the relative motion by the following relations

$$\mathbf{p} = (\mathbf{P} + 2\mathbf{p}_r)/2, \quad \mathbf{p}' = (\mathbf{P} - 2\mathbf{p}_r)/2.$$

Then, as E and E^{Be} are given by

$$E = p_r^2/(2M_I), \quad E^{Be} = (p_r^{Be})^2/(2M_I),$$

the number of Be^8 is

$$\begin{aligned} n_{Be} = n_\alpha^2 (2\pi M_\alpha kT)^{-3} \cdot 16\pi^2 \hbar^3 M_I \int d^3\mathbf{P} \exp[-\mathbf{P}^2/(4M_\alpha kT)] \\ \times \int_0^\infty dp_r [p_r / \{(p_r^2 - (p_r^{Be})^2)^2 + (M_I I^{\alpha Be})^2\}] I_\alpha^{Be} \\ \times \exp[-p_r^2/(M_\alpha kT)]. \end{aligned}$$

The integrand of the last integral has a sharp maximum at $p_r = p_r^{Be}$. Therefore, we can replace p_r by p_r^{Be} in $I_\alpha^{Be} \cdot \exp[-p_r^2/(M_\alpha kT)]$. Performing the integration with this replacement, we obtain

$$\begin{aligned} n_{Be} = 8\pi^3 \hbar^3 (I_\alpha^{Be}/I^{\alpha Be}) n_\alpha^2 (2\pi M_\alpha kT)^{-3} \\ \times \exp[-E^{Be}/kT] \int d^3\mathbf{P} \cdot \exp[-\mathbf{P}^2/(4M_\alpha kT)] \end{aligned} \quad (7a)$$

$$\simeq (4\pi \hbar^2/M_\alpha kT)^{3/2} n_\alpha^2 \exp(-E^{Be}/kT), \quad (I_\alpha^{Be} \sim I^{\alpha Be}, E^{Be} \gg I^{\alpha Be}) \quad (7b)$$

From (7a), the number of Be^8 per unit volume with momenta between \mathbf{P} and $\mathbf{P} + d\mathbf{P}$ is given by

$$8\pi^3 \hbar^3 (I_\alpha^{Be}/I^{\alpha Be}) n_\alpha^2 (2\pi M_\alpha kT)^{-3} \exp(-E^{Be}/kT) \exp(-\mathbf{P}^2/4M_\alpha kT). \quad (8)$$

The rate P_I of the reaction (I), the number of C^{12} produced per unit volume per unit time, is obtained from eqs. (4), (6) and (8),

$$\begin{aligned} P_I = 8\pi^3 \hbar^3 (I_\alpha^{Be}/I^{\alpha Be}) n_\alpha^2 (2\pi M_\alpha kT)^{-3/2} \\ \times \int \int d^3\mathbf{P} d^3\mathbf{p}'' (2E'/M_I')^{1/2} (\pi \hbar^2/2M_I' E') \\ \times [I_\alpha^c I_\gamma^c / \{(E' + E^c)^2 + (I^c/2)^2\}] \exp[-(P^2 + 2p''^2)/(4M_\alpha kT)]. \end{aligned}$$

Taking into account the following experimental facts,

$$E^{Be} \gg I^{\alpha Be}, E^c \gg I^c, I^{Be} \sim I_\alpha^{Be}, I^c = I_\alpha^c + I_\gamma^c, (I_\gamma^c \ll I_\alpha^c),^*)$$

we obtain

$$P_I = 24\sqrt{3} \pi^3 \hbar^5 I_\gamma^c n_\alpha^3 (M_\alpha kT)^{-3} \exp[-(E^{Be} + E^c)/kT]. \quad (9)$$

The resonance energies E^{Be} and E^c are 0.096 Mev and 0.276 Mev, respectively. With these values for E^{Be} and E^c , and with the relation $n_\alpha = \rho X_\alpha/M_\alpha$, the reaction rate is

*) The probability that a C^{12} nucleus excited to the 7.7 Mev state will decay to $\text{Be}^8 + \alpha$ has experimentally been found to be greater than ~ 80 percent⁽⁴⁾.

$$P_I = 1.19 \times 10^{21} \Gamma_{\gamma}^c (10^{-3} \text{ ev}) [\rho (10^5)]^3 X_{\alpha}^3 [T (10^8)]^{-3} \\ \times \exp[-43.2/T (10^8)] \text{ sec}^{-1} \cdot \text{cm}^{-3}, \quad (10)$$

where $T (10^8)$, $\rho (10^5)$ and $\Gamma_{\gamma}^c (10^{-3} \text{ ev})$ mean that the units of temperature T , density ρ and width Γ_{γ}^c are 10^8 K , 10^5 gm/cm^3 and 10^{-3} ev , respectively. As the cascade transition of the second excited state of C^{12} , $0+ \rightarrow 2+ \rightarrow 0+$, is E2 transition, the width Γ_{γ}^c for the first one is of the order of $10^{-3} \sim 10^{-2} \text{ ev}$.

The energy generated by this reaction (I) per unit mass per unit time is

$$\varepsilon_I = \varepsilon_0 [\rho (10^5)]^2 X_{\alpha}^3 [T (10^8)]^{-3} \exp[-43.2/T (10^8)] \text{ erg/gm} \cdot \text{sec}, \quad (11)$$

$$\varepsilon_0 = 1.38 \times 10^{21} \Gamma_{\gamma}^c (10^{-3} \text{ ev}). \quad (12)$$

Our result for the rate of energy production by reaction (I), (II), is different from the one due to Salpeter.³⁾ According to him the rate is

$$\varepsilon_I = 10^3 (\rho/2.5 \times 10^4)^2 [T/(2 \times 10^8 \text{ } ^\circ\text{K})]^{18} X_{\alpha}^3 \text{ erg/gm. sec},$$

near the temperature $2 \times 10^8 \text{ K}$. While, our result is

$$\varepsilon_I = 4.52 \times 10^8 (\rho/2.5 \times 10^4)^2 [T/(2 \times 10^8 \text{ } ^\circ\text{K})]^{18.6} \Gamma_{\gamma}^c (10^{-3} \text{ ev}) X_{\alpha}^3 \text{ erg/gm. sec},$$

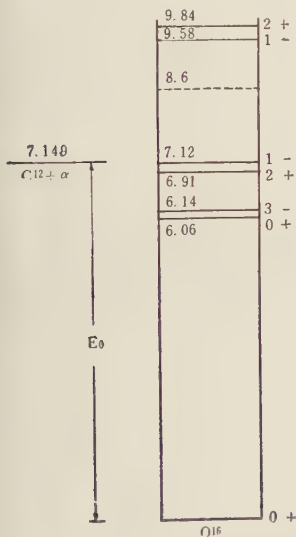


Fig. (2) Energy diagram of the reaction $\text{C}^{12} + \alpha \rightarrow \text{O}^{16} + \gamma$, in Mev.

at the same temperature region (both expressions use the value $\Gamma_{\gamma}^c \sim 10^{-3} \text{ ev}$). Our rate is larger than his by the factor $\sim 10^5$. This difference can be attributed to our adoption of the new experimental data on energy levels.

ii) $\text{C}^{12} + \alpha \rightarrow \text{O}^{16} + \gamma$

We next discuss the second reaction $\text{C}^{12} + \alpha \rightarrow \text{O}^{16} + \gamma$. The energy levels of O^{16} are shown in Fig. (2).⁵⁾ The excited states of O^{16} near the energy E_0 ($E_0 = M_c + M_{\alpha} - M_0$; M_c and M_0 are the masses of C^{12} and O^{16} , respectively.) are 7.12 Mev (1-) state and 8.6 Mev state. The excitation energy of the later state is too high to give an appreciable probability of forming this compound thermonuclear reaction $\text{C}^{12} + \alpha$. Therefore, the state we will consider is the former one, i.e., 7.12-Mev (1-) state.

Assuming that the channel spin is 1, the cross section of the reaction

$$\text{C}^{12} + \alpha \rightarrow \text{O}^{16*} \rightarrow \text{O}^{16} + \gamma$$

is given by

$$\sigma^0(E) = 3\pi (\hbar^2/2M_{11}E) [\Gamma_{\alpha}^0 \Gamma_{\gamma}^0 / \{(E-E^0)^2 + (\Gamma^0/2)^2\}] \\ \simeq 3\pi (\hbar^2/2M_{11}E) [\Gamma_{\alpha}^0 \Gamma_{\gamma}^0 / (E-E^0)^2], \quad (E-E^0 \gg \Gamma^0/2), \quad (13)$$

where

$$M_{II} = M_\alpha M_c / (M_\alpha + M_c) \quad (M_c: \text{mass of } C^{12} \text{ nucleus})$$

$$\Gamma_\alpha^0, \Gamma_\tau^0, \Gamma^0: \text{partial widths of 7.12-Mev state of } O^{16} \text{ for } \alpha\text{-emission, } \gamma\text{-emission and total width of this state,}$$

$$E^0 = 7.12 \text{ Mev} - E_0 < 0$$

The reaction (I) is a resonance reaction. On the contrary, the reaction (II) is a non-resonant one. We should, therefore, take into account the appropriate energy dependence of the widths Γ_α^0 and Γ_τ^0 . They can be expressed as:

$$\begin{aligned} \Gamma_\alpha^0 &= 2 (2M_{II}E/\hbar^2)^{1/2} R_{r_1} \gamma_\alpha, \\ v_1 &= [F_1^2(R) + G_1^2(R)]^{-1}, \end{aligned} \quad (14)$$

where

R : channel radius,

γ_α : reduced width,

F_l, G_l : regular and irregular Coulomb wave functions for $l=1$,

and

$$\Gamma_\tau^0 = A(E_0 + E)^3 \quad (15)$$

The transition $O^{16*} (7.12 \text{ Mev}) \rightarrow O^{16} (\text{ground}) + \gamma$ is E1 transition. A may be estimated from

$$A = \bar{\Gamma}_\tau^0 \times \bar{E}^{-3}, \quad (16)$$

where $\bar{\Gamma}_\tau^0$ is the value of Γ_τ^0 at 7.12 Mev, and $\bar{E} = 7.12 \text{ Mev}$.

We assume that the momentum distributions of both α -particles and C^{12} nuclei are given by the Maxwell law. The rate P_{II} of the reaction (II) is obtained from eqs. (13) (14) and (15),

$$\begin{aligned} P_{II} &= \int \int d^3p d^3p' n_\alpha n_c (M_\alpha M_c)^{-3/2} \times (2\pi kT)^{-3} (2E/M_{II})^{1/2} \\ &\quad \times 3\pi (\hbar^2/2M_{II}E) [2kRv_1 \gamma_\alpha \Gamma_\tau^0 / (E - E^0)^2] \\ &\quad \times \exp[-\{(p^2/M_\alpha) + (p'^2/M_c)\}/2kT] \\ &= 12 \sqrt{\pi} \hbar R n_\alpha n_c \gamma_\alpha A (kT)^{-3/2} M_{II}^{-1} \\ &\quad \times \int E^{1/2} (E_0 + E)^3 (E - E^0)^{-2} (F_1^2 + G_1^2)^{-1} \exp(-E/kT) dE \\ &= 12 \sqrt{\pi} \hbar R (kT)^{-3/2} (M_{II} M_\alpha M_c)^{-1} \gamma_\alpha A^{1/2} X_\alpha X_c \\ &\quad \times \int E^{1/2} (E_0 + E)^3 (E - E^0)^{-2} (F_1^2 + G_1^2)^{-1} \exp(-E/kT) dE, \end{aligned} \quad (17)$$

where n_c is the number of C^{12} per unit volume.

We take for the channel radius of this reaction⁷⁾

$$R = r_0 A^{1/3} + 1.2 \times 10^{-13} \text{ cm}, \quad r_0 = 1.4 \times 10^{-13} \text{ cm}, \quad A = 12.$$

Then, P_{II} is given by

$$P_{II} = 2.83 \times 10^{44} \gamma_{\alpha}(\text{ev}) \bar{I}_{\gamma}^0(\text{ev}) [\rho(10^5)]^2 X_{\alpha} X_c [T(10^8)]^{-3/2} I(T), \quad (18)$$

$$I(T) = \int E'^{1/2} (E_0 + E)^3 (E - E_0)^{-2} (F_1^2 + G_1^2)^{-1} \exp(-E/kT) dE. \quad (19)$$

From this, the energy generation rate ε_{II} by the reaction (II) is given by

$$\varepsilon_{II} = \varepsilon_0' [\rho(10^5)] [T(10^8)]^{-3/2} X_{\alpha} X_c I(T) \text{ erg/gm. sec}, \quad (20)$$

$$\varepsilon_0' = 3.24 \times 10^{34} \gamma_{\alpha}(\text{ev}) \bar{I}_{\gamma}^0(\text{ev}). \quad (21)$$

As the lifetime of the 7.12-Mev state is measured to be $< 8 \times 10^{-15} \text{ sec}^{(9)}$, this transition is inferred to be E1 which is consistent with our assumption. Therefore, \bar{I}_{γ}^0 in eqs. (18)

and (21) may be $1 \sim 10 \text{ ev}$. The sum rule limit⁽⁹⁾ of γ_{α} is $1.6 \times 10^6 \text{ ev}$. Other excited states of O^{16} , of which the lifetimes for α emission are measured, give the values of γ_{α} as about $10^4 \sim 10^5 \text{ ev}$. Therefore, γ_{α} is estimated to be $10^4 \sim 10^5 \text{ ev}$.

Finally, we should calculate the integral $I(T)$. We have calculated it numerically evaluating the Coulomb wave functions at low energies by Abramowitz's method⁽¹⁰⁾. As a table of Bessel-Chriford function is published by Shibagaki⁽¹¹⁾, such a procedure is not too far complicated compared with

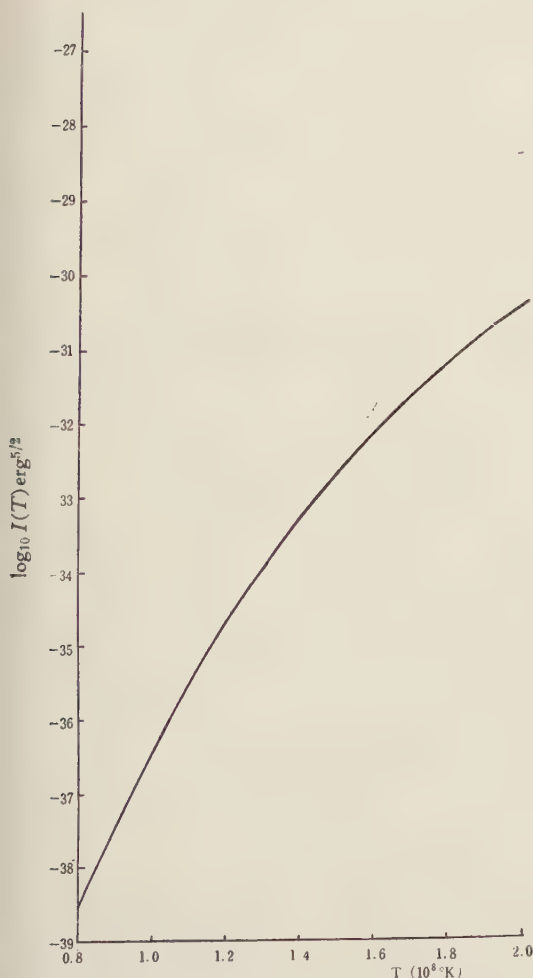


Fig. (3) $I(T)$ (eq. (19)) as a function of temperature T .

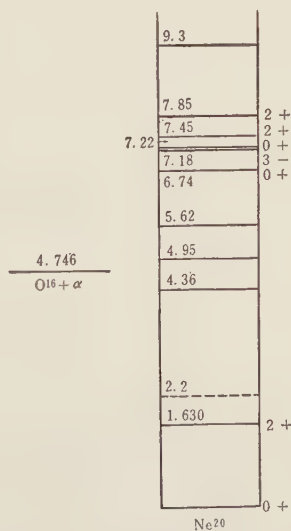


Fig. (4) Energy diagram of the reaction $\text{O}^{16} + \alpha \rightarrow \text{Ne}^{20} + \gamma$, in Mev.

the usual one in which one expands the integrand about the so-called Gamow peak. The former procedure is more accurate than the latter. For high energies we have the values of Coulomb wave functions tabulated by Bloch et al.¹²⁾ The values of $I(T)$ thus obtained for $10^8 < T < 2 \times 10^8$ °K are shown in Fig. (3).

iii) $O^{16} + \alpha \rightarrow Ne^{20} + \gamma$

Experiments on the energy levels of Ne^{20} are scarce. The levels found so far are shown in Fig. (4). The most interesting state for our problem is 4.95-Mev state^{*1}. If it has even spin and even parity, or has odd spin and odd parity, a collision between O^{16} and α can form Ne^{20} in this state. We shall calculate the rate of the reaction (III) assuming that this resonance does occur.

Let channel spin of this reaction be l . Then, the cross section is

$$\sigma^{Ne}(E) = (2l+1) \pi (\hbar^2/2M_{II} E) [I_{\alpha}^{Ne} \Gamma_{\gamma}^{Ne} / \{ (E - E^{Ne})^2 + (\Gamma^{Ne}/2)^2 \}], \quad (22)$$

where

$$M_{II} = M_{\alpha} M_0 / (M_{\alpha} + M_0),$$

E^{Ne} : resonance energy,

Γ_{α}^{Ne} , Γ_{γ}^{Ne} , Γ^{Ne} : partial widths for α -emission, γ -emission and total width of 4.95 Mev state of Ne^{20} .

We assume that the momenta of α -particles and O^{16} nuclei are again given by the Maxwell distribution law. Then, the reaction rate can be obtained by a similar procedure as in i):

$$\begin{aligned} P_{III} &= (2l+1) (2\pi)^{3/2} \hbar^2 n_{\alpha} n_0 (M_{II} kT)^{-3/2} \Gamma_{\alpha}^{Ne} \Gamma_{\gamma}^{Ne} (\Gamma^{Ne})^{-1} \exp(-E^{Ne}/kT) \\ &= (2l+1) (2\pi)^{3/2} \hbar^2 \rho^2 X_{\alpha} X_0 (M_{II} kT)^{-3/2} \Gamma_{\alpha}^{Ne} \Gamma_{\gamma}^{Ne} (\Gamma^{Ne} M_{\alpha} M_0)^{-1} \\ &\quad \times \exp(-E^{Ne}/kT), \end{aligned} \quad (23)$$

$$\begin{aligned} &= 8.00 \times 10^{37} (2l+1) [\rho(10^5)]^2 [T(10^8)]^{-3/2} X_{\alpha} X_0 \times \\ &\quad \times \Gamma_{\alpha}^{Ne}(\text{ev}) \Gamma_{\gamma}^{Ne}(\text{ev}) (\Gamma^{Ne}(\text{ev}))^{-1} \exp[-23.7/T(10^8)] \text{cm}^{-3} \cdot \text{sec}^{-1}. \end{aligned} \quad (24)$$

where n_0 is the number of O^{16} nuclei per unit volume. The rate ε_{III} of energy generation is given by

$$\varepsilon_{III} = \varepsilon_0^{11} \rho(10^5) [T(10^8)]^{-3/2} X_{\alpha} X_0 \exp[-23.7/T(10^8)] \text{erg/gm} \cdot \text{sec}, \quad (25)$$

where

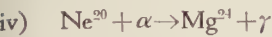
$$\varepsilon_0^{11} = 6.08 \times 10^{27} (2l+1) \Gamma_{\alpha}^{Ne}(\text{ev}) \Gamma_{\gamma}^{Ne}(\text{ev}) / \Gamma^{Ne}(\text{ev}). \quad (26)$$

The situation differs somewhat from that of i), viz., in i) Γ_{α}^{Ne} is far greater than Γ_{γ}^{Ne} , so $\Gamma^{Ne} \sim \Gamma_{\alpha}^{Ne}$ but in iii) Γ_{γ}^{Ne} is far greater than Γ_{α}^{Ne} , thus $\Gamma^{Ne} \sim \Gamma_{\gamma}^{Ne}$. This difference occurs

^{*}1) In the table by Ajzenberg and Lauritsen⁵⁾, one finds a level at 5.4 Mev. Recently, however, Freeman et al.¹³⁾ found that this level should split into two levels at 4.95 Mev and 5.62 Mev.

because the penetration factor for α -emission from Ne^{20} is smaller than that of C^{12} . I_{α}^{Ne} is estimated to be $10^{-15} \sim 10^{-16}$ ev from the penetrability of α -particle based on W. K. B. method.

The numerical results for reactions (I), (II) and (III) are summarized in Table I and Fig. (8) for $I_{\tau}^c=10^{-3}$ ev, $\gamma_{\alpha}=10^5$ ev, $\overline{I}_{\tau}^0=1$ ev and $I_{\alpha}^{Ne}=2.12 \times 10^{-15}$ ev.



The energy diagram of the reaction

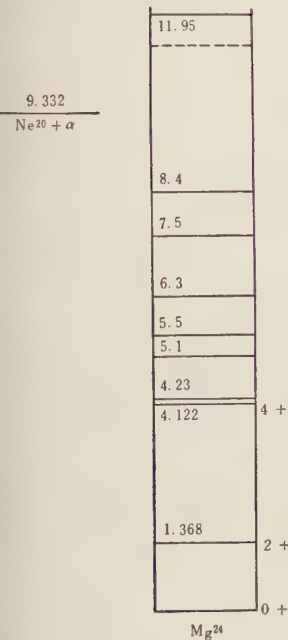
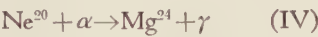


Fig. (5) Energy diagram of the reaction $\text{Ne}^{20} + \alpha \rightarrow \text{Mg}^{24} + \gamma$, in Mev.

is shown in Fig. (5). This figure shows that there has not been observed any excited level of Mg^{24} between 8.4 Mev level and 12.0 Mev level. Both of these levels cannot be reached with an appreciable probability from the entrance channel $\text{Ne}^{20} + \alpha$. But it is hardly concluded in the present stage that in such a wide energy region there are no excited level and that the resonance reaction cannot occur. For the reaction (IV), we can estimate the upper limit of the rate of this reaction, assuming that there is an excited level through which the resonance reaction can occur and which makes the reaction rate maximum at a given temperature.

The rate of this reaction can be obtained by using the same procedure as in iii).

The result is

$$P_{IV} = 6.02 \times 10^{37} (2l+1) [\rho(10^5)]^2 [T(10^8)]^{-3/2} I_{\alpha}^{Mg}(\text{ev}) X_{\alpha} X_{Ne} \times \exp[-116 E^{Mg}(\text{Mev})/T(10^8)] \text{cm}^{-3} \cdot \text{sec}^{-1}, \quad (27)$$

where

l : channel spin,

$I_{\alpha}^{Mg}(\text{ev})$: α -width of the assumed excited level of Mg^{24} ,

$E^{Mg}(\text{Mev})$: resonance energy in Mev.

From eq. (27), the rate of energy production is

$$\epsilon_{IV} = 9.00 \times 10^{27} (2l+1) \rho(10^5) [T(10^8)]^{-3/2} I_{\alpha}^{Mg}(\text{ev}) X_{\alpha} X_{Ne} \times \exp[-116 E^{Mg}(\text{Mev})/T(10^8)] \text{erg}/(\text{gm} \cdot \text{sec}). \quad (28)$$

We have calculated the energy productions for the resonance energies $E^{Mg}(\text{Mev})$ which make the rate P_{IV} of the reaction (IV) maximum at given temperatures. The numerical results are shown in Table I and Fig. (8). From these results it is easily seen that the energy production rate for temperature $< 1.5 \times 10^8$ $^{\circ}\text{K}$ is small enough to

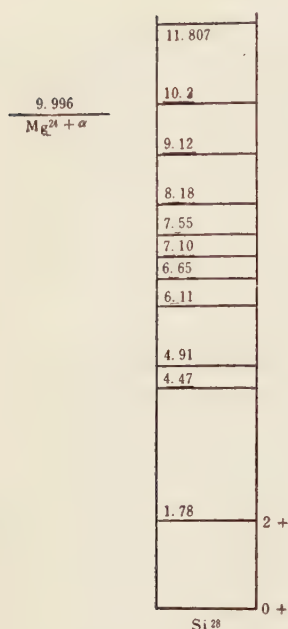


Fig. (6) Energy diagram of the reaction $Mg^{24} + \alpha \rightarrow Si^{28} + \gamma$, in Mev.

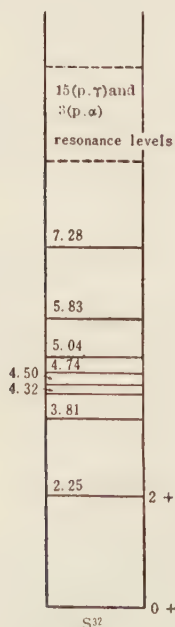
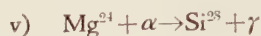


Fig. (7) Energy diagram of the reaction $Si^{28} + \alpha \rightarrow S^{32} + \gamma$, in Mev.

neglect this reaction for any resonance energy.



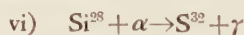
The energy diagram is shown in Fig. (6), including the new levels¹⁴⁾ which have been found recently and one of which, 10.2 Mev level, is the most interesting one for the present problem.

The rate of the reaction (V) and its energy production rate are calculated assuming that the resonance reaction can occur and using completely the same procedure as in iii). The results are

$$P_V = 4.56 \times 10^{37} (2l+1) [\rho(10^5)]^2 [T(10^8)]^{-3/2} \Gamma_{\alpha}^{Si}(\text{ev}) X_{\alpha} X_{Mg} \times \exp[-23.7/T(10^8)] \text{ cm}^{-3} \text{ sec}^{-1}, \quad (29)$$

$$\epsilon_V = 7.30 \times 10^{27} (2l+1) \rho(10^5) [T(10^8)]^{-3/2} \Gamma_{\alpha}^{Si}(\text{ev}) X_{\alpha} X_{Mg} \times \exp[-23.7/T(10^8)] \text{ erg/(gm} \cdot \text{sec)}. \quad (30)$$

The α -width $\Gamma_{\alpha}^{Si}(\text{ev})$ for 10.2 Mev level of Si^{28} is estimated to be 10^{-27} ev by calculating the penetrability of α -particles using the W.K.B. method. The numerical results are shown in Table I and Fig. (8). It is clear from these result that the energy production is very small compared with those from the reactions (I), (II) and (III) and we can neglect this source of energy production in the following discussions.



The energy diagram of the reaction



is shown in Fig. (7). The most interesting level for this problem is 7.28 Mev level¹⁵⁾ of S^{32} . Assuming that the resonance reaction can occur, the reaction rate and the energy production rate of this reaction (VI) are calculated by the same method as in (iii). The results are

$$P_{VI} = 4.00 \times 10^{37} (2l+1) [\rho(10^5)]^2 [T(10^8)]^{-3/2} \Gamma_{\alpha}^S(\text{ev}) X_{\alpha} X_{Si} \times \exp[-39.0/T(10^8)] \text{ cm}^{-3} \cdot \text{sec}^{-1}, \quad (31)$$

$$\epsilon_{VI} = 4.45 \times 10^{27} (2l+1) \rho(10^5) [T(10^8)]^{-3/2} \times \Gamma_{\alpha}^S(\text{ev}) X_{\alpha} X_{Si} \exp[-39.0/T(10^8)] \text{ erg/gm} \cdot \text{sec}. \quad (32)$$

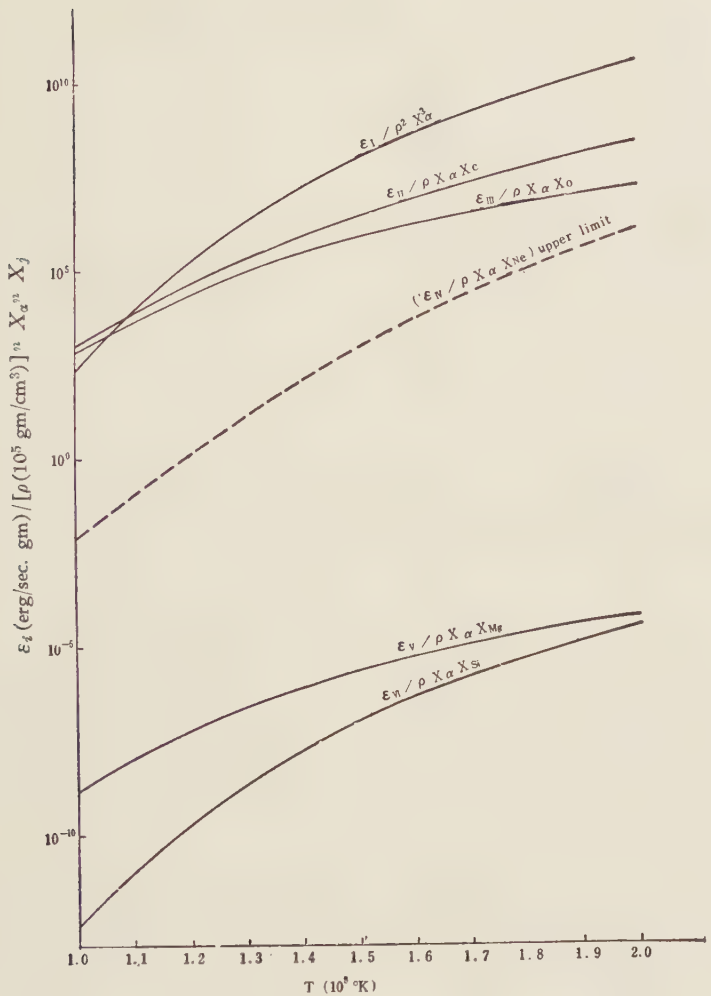


Fig. (8) $\epsilon_i / [\{\rho(10^5)\}^n X_\alpha^n X_j]$ as a function of temperature $T(10^8$ °K), where $\Gamma_{T^0} \sim 10^{-3}$ ev, $\gamma_\alpha \sim 10^5$ ev, $\bar{\Gamma}_{T^0} \sim \text{lev}$, $\Gamma_{\alpha^{Ne}} \sim 2 \times 10^{-15}$ ev, $\Gamma_{\alpha^{Si}} \sim 4 \times 10^{-27}$ ev and $\Gamma_{\alpha^S} \sim 8 \times 10^{-24}$ ev.

Table I. Energy production rates.

| $T_0(10^8 \text{ } ^\circ K)$ | 1.0 | 1.3 | 1.5 |
|---|---------------------------------------|--------------------------------------|--------------------------------------|
| $\epsilon_I / [\rho(10^5)]^2 X_\alpha^3 \alpha$ | $2.46 \times 10^2 (T/T_0)^{40.2}$ | $2.38 \times 10^6 (T/T_0)^{30.2}$ | $1.30 \times 10^8 (T/T_0)^{25.8}$ |
| $\epsilon_{II} / \rho(10^5) X_\alpha X_C^b$ | $1.10 \times 10^3 (T/T_0)^{20.8}$ | $2.41 \times 10^5 (T/T_0)^{18.6}$ | $3.52 \times 10^6 (T/T_0)^{17.5}$ |
| $\epsilon_{III} / \rho(10^5) X_\alpha X_O^c$ | $6.73 \times 10^2 (T/T_0)^{22.2}$ | $1.07 \times 10^5 (T/T_0)^{16.7}$ | $9.82 \times 10^5 (T/T_0)^{14.3}$ |
| $(\epsilon_{IV} / \rho(10^5) X_\alpha X_{Ne})$ upper limit | $8.97 \times 10^{-3} (T/T_0)^{33.3}$ | $1.41 \times 10 (T/T_0)^{34.2}$ | $1.32 \times 10^3 (T/T_0)^{29.4}$ |
| $\epsilon_V / \rho(10^5) X_\alpha X_{Mg}^{d)}$ | $1.61 \times 10^{-9} (T/T_0)^{22.2}$ | $2.56 \times 10^{-7} (T/T_0)^{16.7}$ | $2.34 \times 10^{-6} (T/T_0)^{14.3}$ |
| $\epsilon_{VI} / \rho(10^5) X_\alpha X_{Si}^e)$ | $4.32 \times 10^{-12} (T/T_0)^{37.5}$ | $2.36 \times 10^{-9} (T/T_0)^{28.5}$ | $1.04 \times 10^{-7} (T/T_0)^{24.5}$ |

| $T_0(10^8\text{ }^\circ\text{K})$ | 2.0 | 5.0 | 10.0 |
|---|--------------------------------------|-------------------------------------|-------------------------------------|
| $\varepsilon_I/[\rho(10^5)]^2 X_\alpha^3$ a) | $7.30 \times 10^{10} (T/T_0)^{18.6}$ | | |
| $\varepsilon_{II}/\rho(10^5) X_\alpha X_\beta$ b) | $5.04 \times 10^8 (T/T_0)^{15.9}$ | | |
| $\varepsilon_{III}/\rho(10^5) X_\alpha X_\gamma$ c) | $3.30 \times 10^7 (T/T_0)^{10.3}$ | | |
| $(\varepsilon_{IV}/\rho(10^5) X_\alpha X_{Ne})$ <i>upper limit</i> | $2.42 \times 10^6 (T/T_0)^{27.5}$ | | |
| $\varepsilon_V/\rho(10^5) X_\alpha X_{Mg}$ d) | $7.87 \times 10^{-5} (T/T_0)^{10.3}$ | $2.42 \times 10^{-2} (T/T_0)^{3.2}$ | $9.13 \times 10^{-2} (T/T_0)^{0.9}$ |
| $\varepsilon_{VI}/\rho(10^5) X_\alpha X_{Si}$ e) | $4.48 \times 10^{-5} (T/T_0)^{18.0}$ | $1.37 \times (T/T_0)^{0.3}$ | $2.38 \times 10 (T/T_0)^{2.4}$ |

- a) $\Gamma_T^c = 10^{-5} \text{ ev}$, b) $\gamma_\alpha = 10^5 \text{ ev}$, $\bar{\Gamma}_T^0 = 1 \text{ ev}$, c) $\Gamma_\alpha^{Ne} = 2 \times 10^{-15} \text{ ev}$,
d) $\Gamma_\alpha^{Si} = 4 \times 10^{-27} \text{ ev}$, e) $\Gamma_\alpha^S = 8 \times 10^{-24} \text{ ev}$.

The units of energy production rate and density ρ are erg/(gm·sec) and 10^5 gm/cm^3 .

The α -width $I_\alpha^s(\text{ev})$ for 7.28 Mev level of S^{32} is estimated as 10^{-24} ev . The numerical results are given in Table I and Fig. (8). From these results, we see that this reaction can be neglected in discussing the energy source problem.

§ 3. Energy production by the reaction I, $3\alpha \rightarrow C^{12}$.

It has already been suggested by Öpik²¹⁾ and Salpeter³⁾ that if the temperature and the density of the hydrogen exhausted core are high, the reaction $3\alpha \rightarrow C^{12}$ can supply a large amount of energy. Sandage and Schwarzschild¹⁶⁾ and more recently Hoyle and Schwarzschild¹⁷⁾ have identified some stars of globular cluster as being in this stage of stellar evolution.

In this section we calculate the energy production via the reaction (I), $3\alpha \rightarrow C^{12}$. The energy L_{core} produced in the helium convection core can be easily calculated using the results of section 2, equation (11) and (12). We specify the helium convection core with its central temperature T_c , central density ρ_c and the value of the Emden variable¹⁸⁾ at the outer boundary of the core, x_1 .

$$\begin{aligned} L_{core} &= \int \varepsilon_I 4\pi r^2 \rho \, dr \\ &= 4 \times 10^5 \pi \varepsilon_0 X_\alpha^3 \\ &\quad \times \int [\rho(10^5)]^3 [T(10^8)]^{-3} \exp[-43.2/T(10^8)] r^2 \, dr. \end{aligned} \tag{33}$$

In this convection core, let the ratio of specific heats γ be $5/3$, and introduce the variables x, γ, U and W by the following definitions:

$$\left. \begin{aligned} T(10^8) &= \gamma T_c(10^8), \\ \rho(10^5) &= U \gamma^{5/3} [T_c(10^8)]^{1.5}, \\ x^2 &= W^2 r^2. \end{aligned} \right\} \tag{34}$$

r is the distance from the centre of the star. γ is the Emden function, and its values

for $\gamma=5/3$ are given in tabular form^{18),19)}. U and W are given by

$$\left. \begin{aligned} U &= \rho_c (10^5) / [T_c (10^8)]^{1.5}, \\ W^2 &= 4 \times 10^{-4} \mu \mathcal{R}^{-1} 4\pi G U [T_c (10^8)]^{1.5}, \end{aligned} \right\} \quad (35)$$

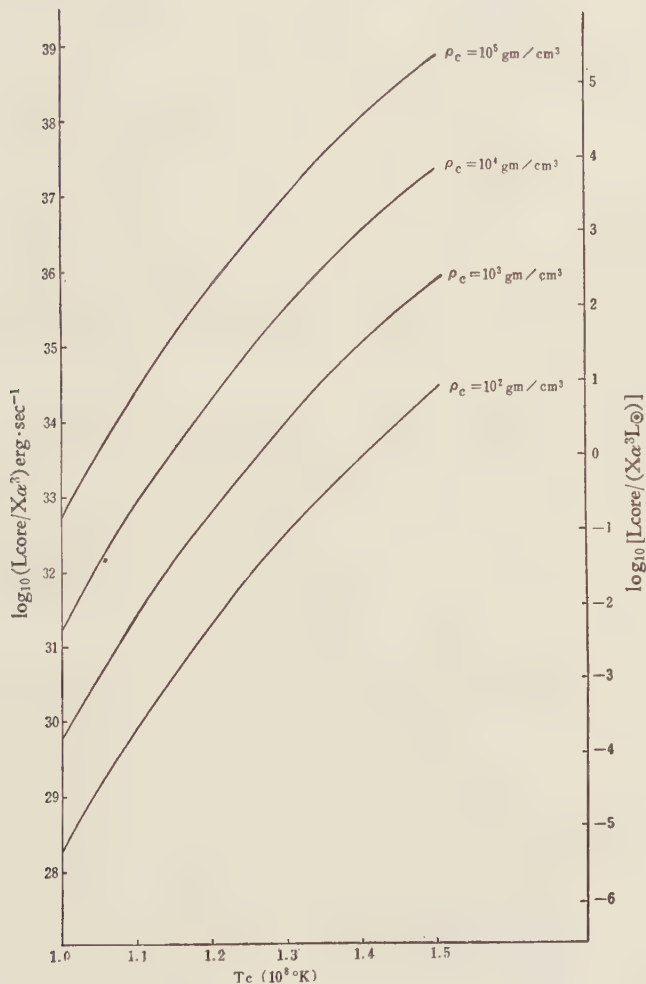


Fig. (9) Energy production as a function of $T_c (10^8 \text{ }^\circ\text{K})$.

where μ is the mean molecular weight, $\mathcal{R} = 8.314 \times 10^7$ erg/deg. mole is the gas constant and $G = 6.670 \times 10^{-8}$ dyne. cm^2/gm^2 is the gravitational constant.

Making use of eqs. (34) and (35),

$$\begin{aligned} L_{\text{core}}/X_\alpha^3 &= 4 \times 10^5 \pi \varepsilon_0 U^3 W^{-3} [T_c (10^8)]^{1.5} \\ &\times \int_0^{x_1} x^2 \gamma^{1.5} \exp [-43.2 / \{T_c (10^8) \gamma\}] dx. \end{aligned} \quad (36)$$

The relation between T_c and $L_{\text{core}}/X_\alpha^3$ is shown in Fig. (9) for four values, 10^2 , 10^3 , 10^4 and 10^5 gm/cm³, of the central density ρ_c . It may be easily seen from this figure that the energy production by this reaction is quite large under the suitable condition of the helium core. Even the total energy of a very luminous star—a star which is several hundred times luminous than the sun—can be produced by this reaction if the central temperature of the helium core is about $1.3 \sim 1.4 \times 10^8$ K and the central density is of the order of $10^3 \sim 10^4$ gm/cm³.

In the study on the stellar models with contracting core, Sandage and Schwarzschild¹⁶⁾ have speculated that the helium reactions may produce the energy at the early phases of the core contraction. But the central temperature reached in their model (for example, $T_c = 1.1 \times 10^8$ K and $\rho_c = 3.6 \times 10^5$ gm/cm³ in model V) was thought to be rather low to produce enough energy. Using our expression eq. (36), the energy produced in the core of their model V is calculated to be $L_{\text{core}} = 1.53 \times 10^{35} X_\alpha^3$ erg/sec and is enough to supply the total energy of stars in the red giant branch of globular clusters. From only this point, the possibility that the energy of such stars is supplied by helium reactions is not completely rejected.

§ 4. Nuclear reactions induced by helium in a helium convection core

The reaction rates for $3\alpha \rightarrow \text{C}^{12}$, $\text{C}^{12} + \alpha \rightarrow \text{O}^{16}$ and $\text{O}^{16} + \alpha \rightarrow \text{Ne}^{20}$ are given by eqs. (9), (17) and (23). For the sake of simplicity we rewrite them as follows:

$$P_I = a T^{-3} n_\alpha^3 \exp [-(E^{B_1} + E^c)/kT], \quad (37)$$

$$P_{II} = b T^{-3/2} n_\alpha n_c I(T), \quad (38)$$

$$P_{III} = c T^{-3/2} n_\alpha n_0 \exp [-E^{Ne}/kT], \quad (39)$$

where

$$a = 24 \sqrt{3} \pi^3 \hbar^5 I_\gamma^c (M_\alpha k)^{-3}, \quad (40)$$

$$b = 12 \sqrt{\pi} \hbar R \gamma_\alpha \bar{I}_\gamma^0 k^{-1/2} M_{II}^{-1} \bar{E}^{-3}, \quad (41)$$

$$c = (2l+1) (2\pi)^{3/2} \hbar^2 I_\alpha^{Ne} (M_{III} k)^{-3/2}, \quad (42)$$

$$M_{II} = M_\alpha M_c / (M_\alpha + M_c), \quad M_{III} = M_\alpha M_0 / (M_\alpha + M_0).$$

and $I(T)$ is an integral given by eq. (17).

To study the reactions (I), (II) and (III) in a helium core, we make the following assumptions:

- i) As these reaction rates depend very much on temperature, a convection core appears in the central region of the helium core, and these reactions occur only in the convection core.
- ii) In this convection core particles are mixed rapidly and the chemical composition within it is uniform.

iii) The central temperature T_c and the central density ρ_c are constant during the time interval with which we are concerned.

iv) The supply of helium from the surroundings into this core and the variation of its volume are negligible.

Let the volume of the convection core be V , and the numbers of α , C^{12} , O^{16} and Ne^{20} in it be N_α , N_c , N_0 and N_{Ne} , respectively. Then,

$$N_\alpha = \int_V n_\alpha dV, \quad N_c = \int_V n_c dV, \quad N_0 = \int_V n_0 dV, \quad N_{Ne} = \int_V n_{Ne} dV. \quad (43)$$

The changing rates of these numbers are

$$\left. \begin{aligned} dN_\alpha/dt &= -3 \int_V P_I dV - \int_V P_{II} dV - \int_V P_{III} dV, \\ dN_c/dt &= \int_V P_I dV - \int_V P_{II} dV, \\ dN_0/dt &= \int_V P_{II} dV - \int_V P_{III} dV, \\ dN_{Ne}/dt &= \int_V P_{III} dV. \end{aligned} \right\} \quad (44)$$

From eq. (43) and $n_i = \rho X_i/M_i$, these differential equations can be written as follows:

$$\left. \begin{aligned} dX_\alpha/dt &= -(\alpha_1 X_\alpha^2 + \alpha_2 X_c + \alpha_3 X_0) X_\alpha, \\ dX_c/dt &= (\beta_1 X_\alpha^2 - \beta_2 X_c) X_\alpha, \\ dX_0/dt &= (\gamma_1 X_c - \gamma_2 X_0) X_\alpha, \\ dX_{Ne}/dt &= \delta X_0 X_\alpha, \end{aligned} \right\} \quad (45)$$

where

$$\left. \begin{aligned} \alpha_1 &= 3aJ_1/(M_\alpha^2 J_0) = 2.37 \times 10^7 \Gamma_\tau^c(\text{ev}) J_1/J_0, \\ \alpha_2 &= bJ_2/(M_c J_0) = 1.88 \times 10^{16} \gamma_\alpha(\text{ev}) \bar{\Gamma}_\tau^0(\text{ev}) J_2/J_0, \\ \alpha_3 &= cJ_3/(M_0 J_0) = 5.31 \times 10^9 (2l+1) \Gamma_\alpha^{Ne}(\text{ev}) J_3/J_0, \\ \beta_1 &= aM_c J_1/(M_\alpha^3 J_0) = 2.37 \times 10^7 \Gamma_\tau^c(\text{ev}) J_1/J_0, \\ \beta_2 &= bJ_2/(M_\alpha J_0) = 5.64 \times 10^{16} \gamma_\alpha(\text{ev}) \bar{\Gamma}_\tau^0(\text{ev}) J_2/J_0, \\ \gamma_1 &= bM_0 J_2/(M_\alpha M_c J_0) = 7.52 \times 10^{16} \gamma_\alpha(\text{ev}) \bar{\Gamma}_\tau^0(\text{ev}) J_2/J_0, \\ \gamma_2 &= cJ_3/(M_\alpha J_0) = 2.12 \times 10^{10} (2l+1) \Gamma_\alpha^{Ne}(\text{ev}) J_3/J_0, \\ \delta &= cM_{Ne} J_3/(M_\alpha M_0 J_0) = 2.65 \times 10^{10} (2l+1) \Gamma_\alpha^{Ne}(\text{ev}) J_3/J_0, \end{aligned} \right\} \quad (46)$$

$$\left. \begin{aligned} J_0 &= \int_V \rho dV, \\ J_1 &= \int_V T^{-3} \rho^3 \exp[-(E^{Be} + E^c)/kT] dV, \\ J_2 &= \int_V T^{-3/2} \rho^2 I(T) dV, \\ J_3 &= \int_V T^{-3/2} \rho^2 \exp[-E^{Ne}/kT] dV. \end{aligned} \right\} \quad (47)$$

For the sake of simplicity, T and ρ are used instead of $T(10^8)$ and $\rho(10^5)$ in eq. (47). In the following we shall use this abbreviation. The integrals (47) can be calculated numerically, when T , ρ and V are given.

From eqs. (34) and (35), the integrals (47) become as follows:

$$\left. \begin{aligned} J_1 &= 4\pi U^3 W^{-3} T_c^{1.5} \int_0^{x_1} \gamma^{1.5} \exp(-43.2/\gamma T_c) x^2 dx, \\ J_2 &= 4\pi U^2 W^{-3} T_c^{1.5} \int_0^{x_1} \gamma^{1.5} I(\gamma T_c) x^2 dx, \\ J_3 &= 4\pi U^2 W^{-3} T_c^{1.5} \int_0^{x_1} \gamma^{1.5} \exp(-23.7/\gamma T_c) x^2 dx, \\ J_0 &= 4\pi U W^{-3} T_c^{1.5} \int_0^{x_1} \gamma^{1.5} x^2 dx. \end{aligned} \right\} \quad (48)$$

These integrals can be evaluated once T_c , ρ_c and x_1 are known. As for these values we shall discuss in the following section.

§ 5. Variation of the chemical compositions and energy production by helium capture reactions

Once the helium in the hydrogen exhausted core begins to be converted into carbon, oxygen and neon will be produced by successive capture of remaining helium by reaction (II) and (III). Thus the chemical composition in the core will be altered according to the differential equations (45). Then the energy produced in the core is not only supplied by the reaction (I), but also supplied by reactions (II) and (III).

In order to know the variation of the luminosity of a star whose energy is supplied by these reactions, we must know the variation of the chemical composition, X_α , X_c , X_0 , and X_N , besides the variation of the central temperature and the central density in connection with the stellar structure.

In this section we restrict ourselves to the case in which the central temperature, T_c , the central density ρ_c and the outer boundary x_1 of the helium convection core have constant values according to the assumptions iii) and iv). As a set of these values which characterise the central helium convection core, we use the one calculated by one of the authors (Obi) for an inhomogeneous stellar model with 1.3 solar masses²⁰⁾. Assuming the energy is produced both by CN-cycle in a thin shell surrounding the helium core and by the reaction (I) in the central helium convection core, this model is applied for stars in the horizontal branch of the globular cluster. The adopted values are

$$\left. \begin{aligned} T_c &= 1.332 \times 10^8 \text{ }^\circ\text{K}, \\ \rho_c &= 6.992 \times 10^3 \text{ gm/cm}^3, \\ x_1 &= 1.57856. \end{aligned} \right\} \quad (49)$$

Equations (45) have been solved then by numerical quadrature using these values. As to the widths l_{τ}^c , γ_α , etc., appearing in eqs. (46), we can only estimate the order of magnitudes of their values, so the unique set of solutions cannot be obtained. Defining the following two parameters,

$$\xi = \gamma_\alpha \bar{l}_{\tau}^0 / l_{\tau}^c, \text{ and } \eta = (2l+1) l_{\alpha}^{Ne} / l_{\tau}^c, \quad (50)$$

the solutions are obtained for some plausible values for these two parameters. Equations (45) contain time t as an independent variable. But it is convenient to change the independent variable t to X_α :

$$\left. \begin{aligned} dX_c/dX_\alpha &= -(\beta_1 X_\alpha^2 - \beta_2 X_c) / (\alpha_1 X_\alpha^2 + \alpha_2 X_c + \alpha_3 X_0), \\ dX_0/dX_\alpha &= -(\gamma_1 X_c - \gamma_2 X_0) / (\alpha_1 X_\alpha^2 + \alpha_2 X_c + \alpha_3 X_0), \\ dX_{Ne}/dX_\alpha &= -\delta X_0 / (\alpha_1 X_\alpha^2 + \alpha_2 X_c + \alpha_3 X_0), \\ dt/dX_\alpha &= -1 / [(\alpha_1 X_\alpha^2 + \alpha_2 X_c + \alpha_3 X_0) X_\alpha]. \end{aligned} \right\} \quad (51)$$

The initial conditions of these eqs. are taken to be

$$\left. \begin{aligned} X_\alpha &= 1, \\ X_c &= X_0 = X_{Ne} = 0, \\ t &= 0. \end{aligned} \right\} \quad (52)$$

The results of our numerical calculation with different values of parameters ξ and η are shown in Figs. (10) ~ (19). The comparison of two figures having the same ξ and different η shows that there is a small difference in X_c and X_0 between two cases. Furthermore, it is seen in Figs. (15) ~ (19) that X_α shows almost the same behavior, whichever set of the constants we may choose. If the sets of the constants ξ and η are such as the ones that appear in Figs. (11), (16), (13) and (18), the rate of the reaction $O^{16} + \alpha \rightarrow Ne^{20} + \gamma$ should be large enough compared with the rate of $C^{12} + \alpha \rightarrow O^{16} + \gamma$, that is, the reaction $C^{12} + \alpha \rightarrow O^{16} + \gamma$ is immediately followed by $O^{16} + \alpha \rightarrow Ne^{20} + \gamma$. On the other hand, if $\xi = 10^0$ and $\eta = 2.12 \times 10^{-12}$ as in Figs. (14) and (19), the reaction $3\alpha \rightarrow C^{12} + \gamma$ is immediately followed by $C^{12} + \alpha \rightarrow O^{16} + \gamma$.

To get X_c , X_0 and X_{Ne} as functions of X_α , only the ratios ξ and η between reaction widths are needed. But to get them as functions of t , we must know the magnitudes of reaction widths, I_α^c , γ_α and I_α^{Ne} . In Figs. (15) ~ (19), I_T^o is taken to be 10^{-3} ev. If I_T^c be f times 10^{-3} ev, t should be replaced by $1/f$ times the one which appears in these figures.

As the chemical composition of the central convection core changes, the energy production within the core will be altered. The rates of energy generation by the reactions (I), (II) and (III) are given by (11), (20) and (25). Making use of these results, the nuclear energy production within the core is given by

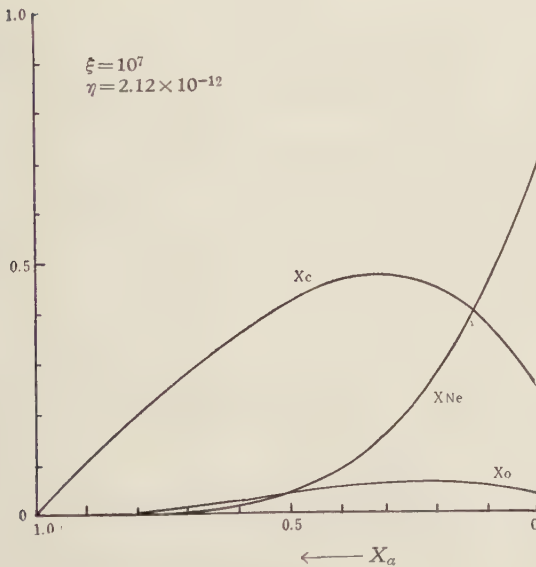


Fig. (10) Concentration (by weight) X_c , X_0 and X_{Ne} of C^{12} , O^{16} and Ne^{20} as functions of X_α with constants $\xi = 10^7$, and $\eta = 2.12 \times 10^{-12}$.

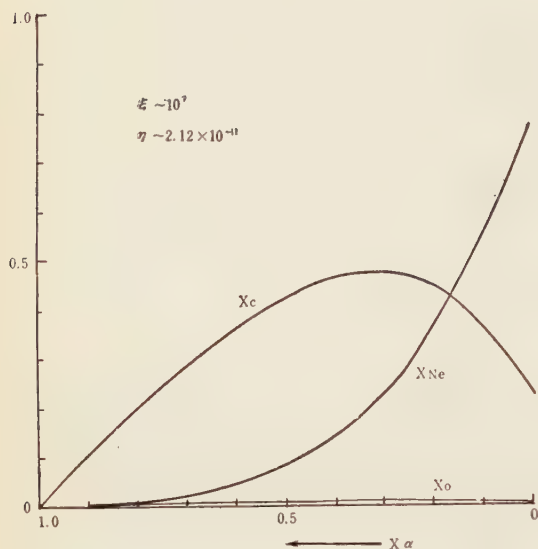


Fig. (11) Concentration (by weight) X_c , X_0 , and X_{Ne} of Cl^{32} , O^{16} and Ne^{20} as functions of X_α with constants $\xi=10^7$, and $\eta=2.12 \times 10^{-11}$.

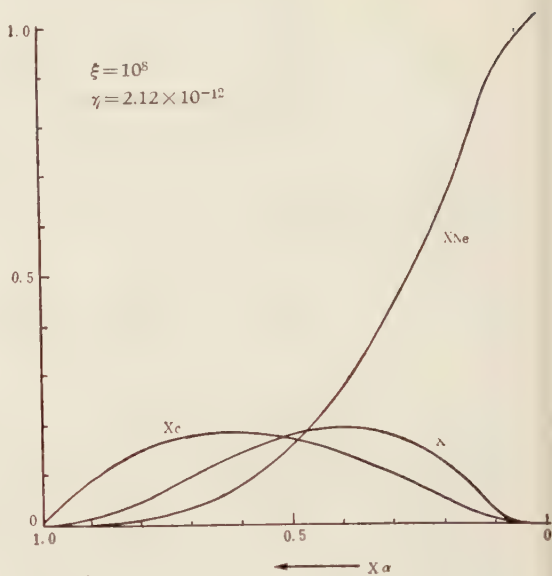


Fig. (12) Concentration (by weight) X_c , X_0 and X_{Ne} of Cl^{32} , O^{16} and Ne^{20} as functions of X_α with constants $\xi=10^8$, and $\eta=2.12 \times 10^{-12}$.

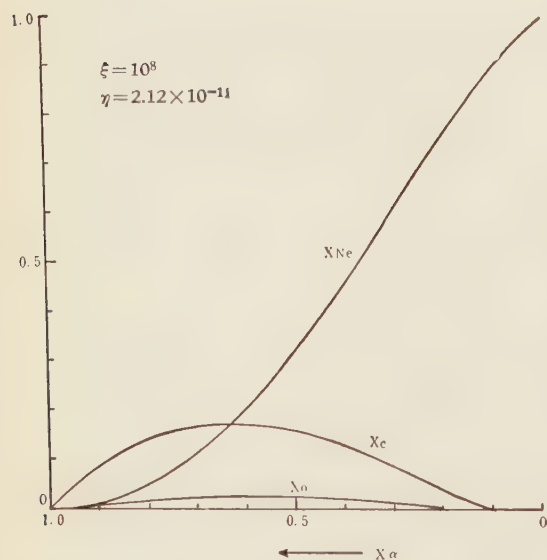


Fig. (13) Concentration (by weight) X_c , X_0 and X_{Ne} of Cl^{32} , O^{16} and Ne^{20} as functions of X_α with constants $\xi=10^8$ and $\eta=2.12 \times 10^{-11}$.

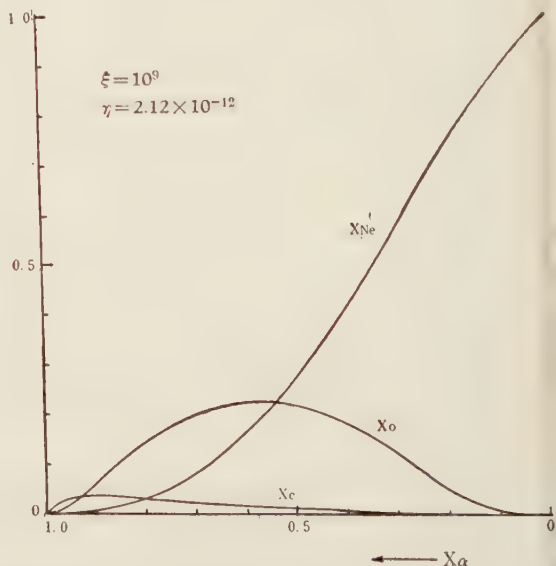


Fig. (14) Concentration (by weight) X_c , X_0 and X_{Ne} of Cl^{32} , O^{16} and Ne^{20} as functions of X_α with constants $\xi=10^9$, and $\eta=2.12 \times 10^{-12}$.

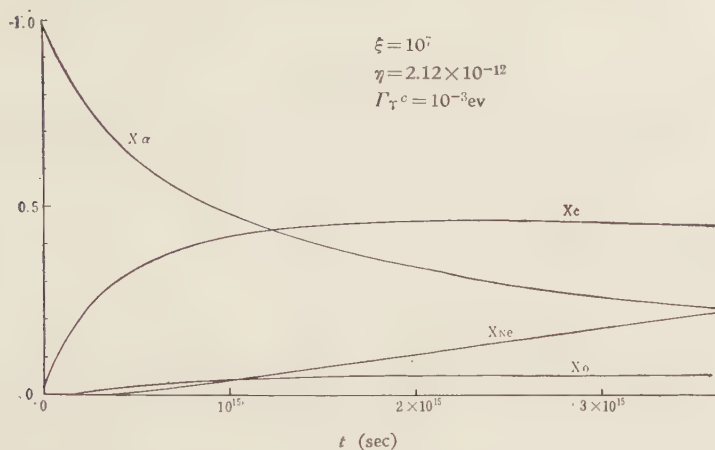


Fig. (15) Concentration (by weight) X_α , X_c , X_o and X_{Ne} of α , C^{12} , O^{16} and Ne^{20} as functions of time t with $\xi = 10^7$, $\eta = 2.12 \times 10^{-12}$ and $\Gamma_{Tc} = 10^{-3} \text{ ev}$.

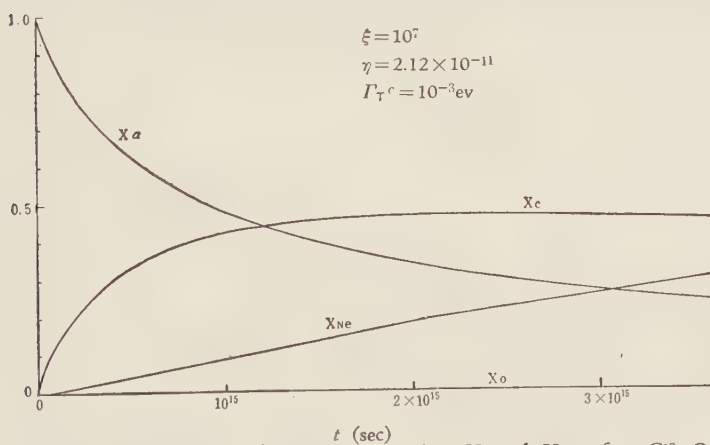


Fig. (16) Concentration (by weight) X_α , X_c , X_o and X_{Ne} of α , C^{12} , O^{16} and Ne^{20} as functions of time t with $\xi = 10^7$, $\eta = 2.12 \times 10^{-11}$ and $\Gamma_{Tc} = 10^{-3} \text{ ev}$.

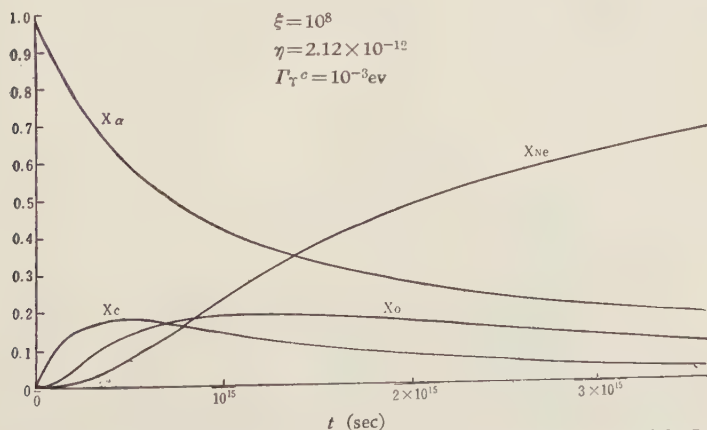


Fig. (17) Concentration (by weight) X_α , X_c , X_o and X_{Ne} of α , C^{12} , O^{16} and Ne^{20} as functions of time t with $\xi = 10^8$, $\eta = 2.12 \times 10^{-12}$ and $\Gamma_{Tc} = 10^{-3} \text{ ev}$.

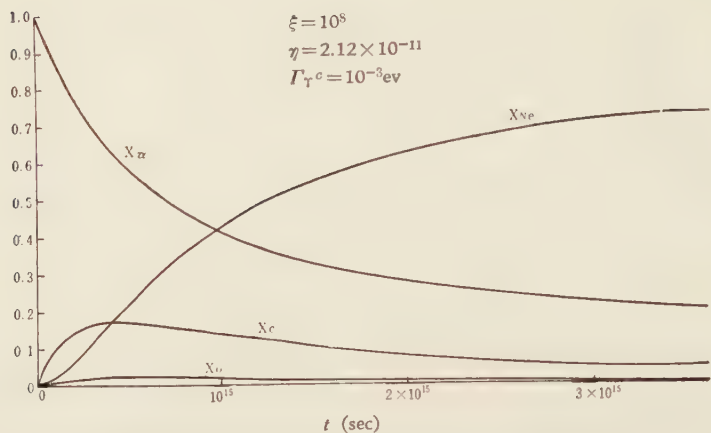


Fig. (18) Concentration (by weight) X_α , X_c , X_o and X_{Ne} of α , C^{12} , O^{16} and Ne^{20} as functions of time t with $\xi = 10^8$, $\eta = 2.12 \times 10^{-11}$ and $\Gamma_T^c = 10^{-3}$ ev.

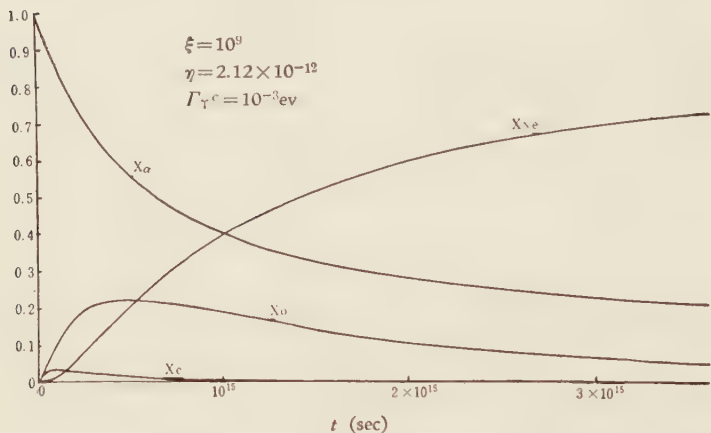


Fig. (19) Concentration (by weight) X_α , X_c , X_o and X_{Ne} of α , C^{12} , O^{16} and Ne^{20} as functions of time t with $\xi = 10^9$, $\eta = 2.12 \times 10^{-12}$ and $\Gamma_T^c = 10^{-3}$ ev

$$L_{core} = 4\pi \int \epsilon \rho r^2 dr, \tag{53}$$

$$\epsilon = \epsilon_I + \epsilon_{II} + \epsilon_{III}. \tag{54}$$

Transforming the above equation by (34), we obtain

$$\begin{aligned} L_{core} = & 4\pi \times 10^5 U^2 W^{-3} \\ & \times [\epsilon_0 U X_\alpha^3 \int_0^{x_1} \gamma^{1.5} x^2 \exp[-43.2/\gamma T_c] dx \\ & + \epsilon_0' X_\alpha X_c \int_0^{x_1} \gamma^{1.5} I'(y) x^2 dx \\ & + \epsilon_0'' X_\alpha X_o \int_0^{x_1} \gamma^{1.5} x^2 \exp(-23.7/\gamma T_c) dX], \end{aligned} \tag{55}$$

where

$$I'(\gamma) = \int E^{1/2} (E_0 + E)^{-3} (E - E_0)^{-2} (G_1^2 + F_1^2)^{-1} \times \exp(-E/k\gamma T_c) dE. \quad (56)$$

$I'(\gamma)$ is calculated in § 2, ii). Using the values of U , W and $I'(\gamma)$, one finds

$$L_{core} = 4.11 \times 10^{31} \mu^{-3/2} [1.65 \times 10^4 P_{\tau}^c (10^{-3} \text{ ev}) X_{\alpha}^2 + 3.56 \times 10^{-1} \gamma_{\alpha}(\text{ev}) \bar{P}_{\tau}^0(\text{ev}) X_c + 8.13 \times 10^{18} (2l+1) P_{\alpha}^{N_2}(\text{ev}) X_0] X_{\alpha}. \quad (57)$$

Where μ is the mean molecular weight of the central convection core and is given by

$$\mu = [(3/4) X_x + (7/12) X_c + (9/16) X_0 + (11/20) X_{Ne}]^{-1} \quad (58)$$

μ is also obtained numerically. The nuclear energy production within the core thus obtained is shown in Figs. (20) and (21).

It is seen in these figures, that energy production gradually decreases, as time passes and X_{α} decreases, for any set of constants ξ and η . Here, P_{τ}^c is taken to be $= 10^{-3} \text{ ev}$. If P_{τ}^c be f times 10^{-3} ev , L_{core} in Figs. (20) and (21) should be f times the one which

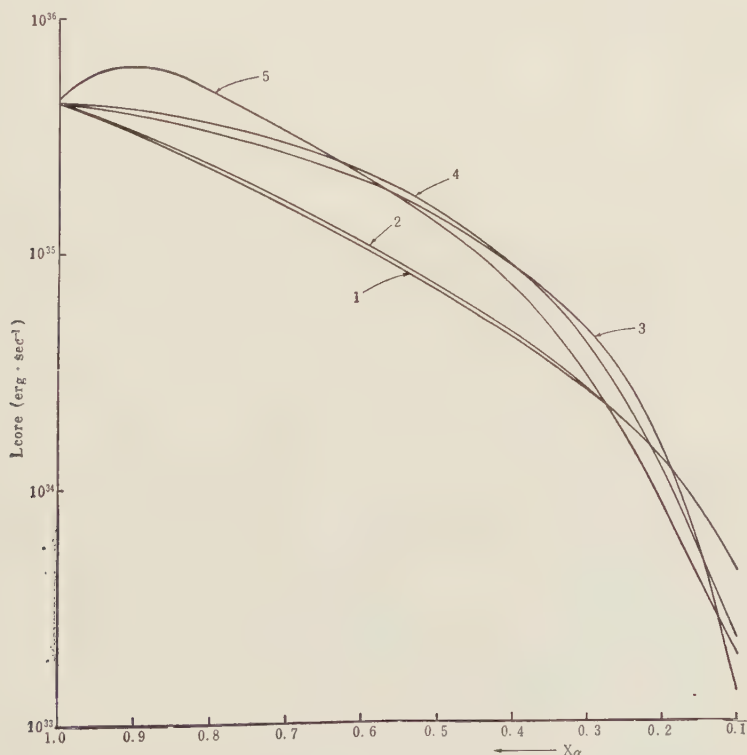


Fig. (20) Energy production L_{core} as a function of X_{α} , with $P_{\tau}^c = 10^{-3} \text{ ev}$. Each curve corresponds to one of the sets of ξ and η , 1: $\xi = 10^7$, $\eta = 2.12 \times 10^{-12}$, 2: $\xi = 10^7$, $\eta = 2.12 \times 10^{-11}$, 3: $\xi = 10^8$, $\eta = 2.12 \times 10^{-12}$, 4: $\xi = 10^8$, $\eta = 2.12 \times 10^{-11}$, 5: $\xi = 10^9$, $\eta = 2.12 \times 10^{-12}$.

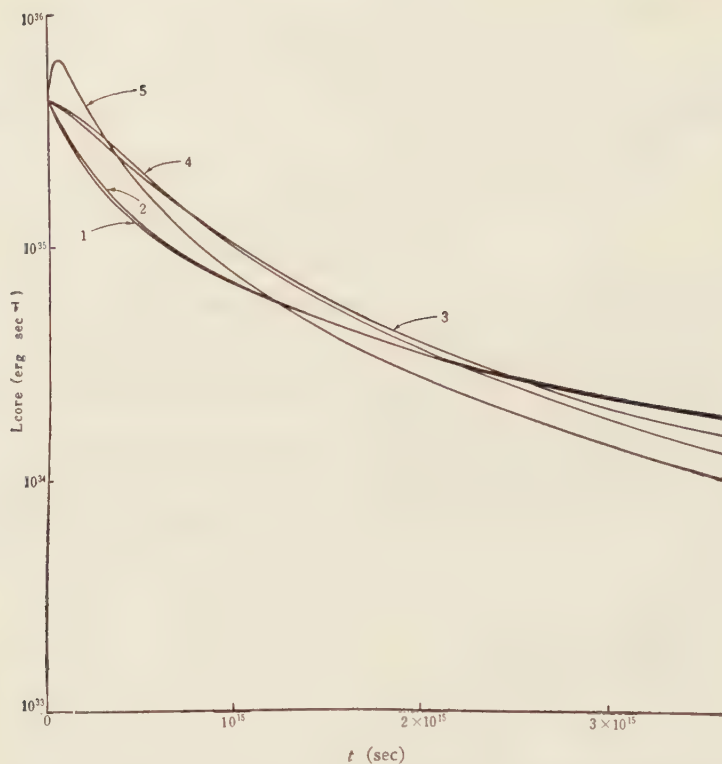


Fig. (21) Energy production L_{core} as a function of the time t with $\Gamma_T^0 = 10^{-3}$ ev. Each curve corresponds to one of the sets of ξ and η , 1: $\xi = 10^7$, $\eta = 2.12 \times 10^{-12}$, 2: $\xi = 10^7$, $\eta = 2.12 \times 10^{-11}$, 3: $\xi = 10^8$, $\eta = 2.12 \times 10^{-12}$, 4: $\xi = 10^8$, $\eta = 2.12 \times 10^{-11}$, 5: $\xi = 10^9$, $\eta = 1.12 \times 10^{-12}$.

is given in these figures, and t should become $1/f$ times the one in Fig. (21).

§ 6. Nuclear reactions among C^{12} , O^{16} , Ne^{20}

As discussed in the preceding sections, the α -capture reactions produce C^{12} , O^{16} , Ne^{20} , etc. In the present section we consider whether the nuclear reactions among these nuclei themselves can effectively occur or not in actual stars. According to the data summarized in the articles by Ajzenberg and Lauritsen⁵¹⁾, and by Endt and Kluwyer⁵²⁾ we find the following energetically possible reactions using the values of mass defects.

The reaction rate for the formation of compound state by colliding two nuclei with each other will be given as follows.

$$P = (2\pi)^{3/2} \hbar^2 \rho^2 X_A X_B (kTM)^{-3/2} (M_A M_B)^{-1} P_1 (2J+1) \times \exp(-116 E/kT) \text{ cm}^{-3} \text{ sec}^{-1}, \quad (59)$$

where E : an energy value of relative motion between two nuclei corresponding to the resonance level, J : the channel spin, M_A and M_B : masses of A and B nuclei, $M = M_A M_B / (M_A + M_B)$: the reduced mass, P_1 : reaction width in the process $A+B \rightarrow$ the com-

pound state of A and B . Expressing M_A and M_B in mass unit, P_1 and E in Mev, ρ in 10^5 gr/cm³ and T in 10^8 °K, P is reduced to

Table II. Energetically possible nuclear reactions among C¹², O¹⁶ and Ne²⁰.

| | | |
|---------------------------------------|---|---|
| C ¹² + C ¹² → | $\left\{ \begin{array}{l} \text{Mg}^{24} + \gamma + 13.95 \text{ Mev} \\ \text{Na}^{23} + \text{p} + 2.25 \text{ Mev} \\ \text{Ne}^{20} + \alpha + 4.62 \text{ Mev} \end{array} \right.$ | negligible |
| C ¹² + O ¹⁶ → | $\left\{ \begin{array}{l} \text{Si}^{28} + \gamma + 16.8 \text{ Mev} \\ \text{Al}^{27} + \text{p} + 5.21 \text{ Mev} \\ \text{Mg}^{24} + \alpha + 6.80 \text{ Mev} \end{array} \right.$ | negligible |
| C ¹² + Ne ²⁰ → | $\left\{ \begin{array}{l} \text{S}^{32} + \gamma + 19.0 \text{ Mev} \\ \text{S}^{31} (\rightarrow \text{P}^{31} + \beta^+ : 2.4 \text{ sec.}) \\ \quad + \text{n} + \gamma + 4.2 \text{ Mev} \\ \text{P}^{31} + \text{p} + 10.1 \text{ Mev} \\ \text{Si}^{28} + \alpha + 12.0 \text{ Mev} \\ \text{Mg}^{24} + \text{Be}^8 + 2.0 \text{ Mev} \\ \quad (\text{Be}^8 \rightarrow 2\alpha : < 4 \times 10^{-15} \text{ sec.}) \\ \text{O}^{16} + \text{O}^{16} + 2.4 \text{ Mev} \end{array} \right.$ | <p>probably negligible</p> <p>negligible</p> <p>negligible</p> |
| O ¹⁶ + O ¹⁶ → | $\left\{ \begin{array}{l} \text{S}^{32} + \gamma + 16.6 \text{ Mev} \\ \text{S}^{31} (\rightarrow \text{P}^{31} + \beta^+ : 2.4 \text{ sec.}) \\ \quad + \text{n} + 1.8 \text{ Mev} \\ \text{P}^{31} + \text{p} + 7.7 \text{ Mev} \\ \text{Si}^{28} + \alpha + 9.7 \text{ Mev} \end{array} \right.$ | probably negligible |
| O ¹⁶ + Ne ²⁰ → | $\left\{ \begin{array}{l} \text{A}^{36} + \gamma + 18.5 \text{ Mev} \\ \text{A}^{35} (\rightarrow \text{Cl}^{35} + \beta^+ : 1.9 \text{ sec.}) \\ \quad + \text{n} + 3.8 \text{ Mev} \\ \text{Cl}^{35} + \text{p} + 10.0 \text{ Mev} \\ \text{S}^{32} + \alpha + 11.8 \text{ Mev} \\ \text{P}^{31} + \text{Li}^5 + 1.2 \text{ Mev} \\ \text{Si}^{28} + \text{Be}^8 + 4.8 \text{ Mev} \\ \quad (\text{Be}^8 \rightarrow 2\alpha : < 4 \times 10^{-15} \text{ sec.}) \\ \text{Mg}^{24} + \text{C}^{12} + 2.2 \text{ Mev} \end{array} \right.$ | <p>probably negligible</p> <p>probably small</p> <p>negligible</p> <p>negligible</p> <p>negligible</p> |
| Ne ²⁰ + Ne ²⁰ → | $\left\{ \begin{array}{l} \text{Ca}^{40} + \gamma + 20.9 \text{ Mev} \\ \text{Ca}^{39} (\rightarrow \text{K}^{39} + \beta^+ : 0.9 \text{ sec.}) \\ \quad + \text{n} + 5.0 \text{ Mev} \\ \text{K}^{39} + \text{p} + 12.5 \text{ Mev} \\ \text{K}^{38} (\rightarrow \text{A}^{38} + \beta^+ : 7.7 \text{ min.}) \\ \quad + \text{d} + 1.5 \text{ Mev} \\ \text{A}^{37} (\rightarrow \text{Cl}^{37} \text{ K. L. capture} \\ \quad : 34 \text{ d}) + \text{He}^3 + 2.0 \text{ Mev} \\ \text{A}^{36} + \alpha + 13.7 \text{ Mev} \\ \text{Cl}^{35} + \text{Li}^5 + 3.4 \text{ Mev} \\ \text{S}^{32} + \text{Be}^8 + 7.0 \text{ Mev} \\ \quad (\text{Be}^8 \rightarrow 2\alpha : < 4 \times 10^{-15} \text{ sec.}) \\ \text{Si}^{28} + \text{C}^{12} + 7.4 \text{ Mev} \\ \text{Mg}^{24} + \text{O}^{16} + 4.6 \text{ Mev} \end{array} \right.$ | <p>probably negligible</p> <p>probably main</p> <p>negligible</p> <p>negligible</p> <p>negligible</p> <p>negligible</p> <p>negligible</p> |

$$P = 1.02 \times 10^{48} \rho^2 X_A X_B (MT)^{-3/2} (M_A M_B)^{-1} \times P_1 (2J+1) \exp(-E/T). \quad (60)$$

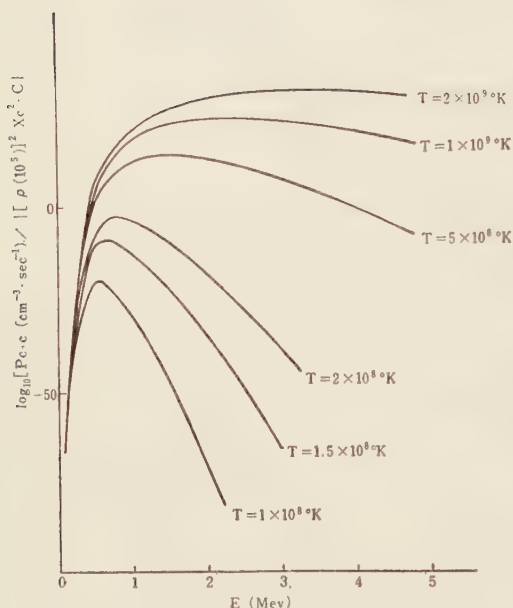


Fig. (22) $P_{c+c}/(\rho^2 X_c^2 C)$ as a function of $E(\text{Mev})$.

In order to examine to what extent this reaction competes with the reaction $C^{12} + \alpha \rightarrow O^{16*}$, we calculate the ratio of both reaction rates at the most favourable value of E for the reaction $C^{12} + C^{12} \rightarrow Mg^{24*}$,

$$\frac{P_{c+c}}{P_{c+\alpha}} = \begin{cases} 2.55 \times 10^{-32} (C/\gamma_\alpha \bar{P}_T^0) (X_c/X_\alpha) : T=1(10^8 \text{ }^\circ\text{K}) \\ 1.60 \times 10^{-21} (C/\gamma_\alpha \bar{P}_T^0) (X_c/X_\alpha) : T=2 \\ 1.82 \times 10^{-15} (C/\gamma_\alpha \bar{P}_T^0) (X_c/X_\alpha) : T=5. \end{cases}$$

If we assume that $C \sim 1$, $\gamma_\alpha \sim 10^4$, and $\bar{P}_T^0 \sim 1$, $P_{c+c}/P_{c+\alpha}$ is very small even for a comparatively high temperature: $T=5 \times 10^8$ °K. Thus the reaction $C^{12} + C^{12} \rightarrow Mg^{24*}$ does not proceed unless α -nucleus is consumed entirely or the temperature becomes high enough. We have assumed here that the reactions are due to only one resonance level, but the level interval will be $20 \sim 100$ Kev in the energy region considered here, so the reaction rate could be larger than our estimation by a factor ~ 10 .

Next we will consider the other reactions among C^{12} , O^{16} and Ne^{20} . The rate of reaction $A+B \rightarrow (AB)$, per one nucleus A, is summarized in Table III, assuming that the density of nucleus B is $\rho X_B = 10^{21} M_H \text{ gr/cm}^3$, and that the level spacing is 100 kev. From the table we see that,

- 1) the reaction between C^{12} nuclei becomes effective at such high temperature as

P_1 appeared in (59) and (60) is calculated in W.K.B approximation as the following way

$$\Gamma_1 = C \exp[-2f(E)],$$

$$f(E) = \int_{R_0}^{r_0} \kappa_l(r) dr,$$

$$\kappa_l^2(r) = l(l+1)/r^2 + 2MZZ'e^2/(\hbar r) - 2ME/\hbar,$$

$$\kappa_l(r_0) = 0,$$

where C is a quantity which is of the order of magnitude $2/(2\pi)$ times the level spacing of the compound states with same J and same parity. R is taken to be $1.4 \times 10^{-13} (A^{1/3} + B^{1/3}) \text{ cm}$.

First we consider the reaction $C^{12} + C^{12} \rightarrow Mg^{24*}$. The reaction rate for this process is shown in Fig. (22) as a function of T and of E .

Table III. Reaction rates between two nuclei among C^{12} , O^{16} and Ne^{20} in year $^{-1}$

| T | E | $C^{12} + C^{12}$ | $O^{16} + O^{16}$ | $Ne^{20} + Ne^{20}$ | $C^{12} + Ne^{20}$ |
|-----------------|-------|-------------------|-------------------|---------------------|--------------------|
| 10^9 °K | 1 Mev | $10^{-8.4}$ | $10^{-37.6}$ | $10^{-84.0}$ | $10^{-32.7}$ |
| | 2 | $10^{-2.9}$ | $10^{-20.7}$ | $10^{-49.6}$ | $10^{-17.4}$ |
| | 3 | $10^{-3.5}$ | $10^{-16.0}$ | $10^{-37.4}$ | $10^{-13.6}$ |
| 5×10^8 | 1 | $10^{-12.9}$ | $10^{-42.7}$ | $10^{-88.6}$ | $10^{-37.3}$ |
| | 2 | $10^{-12.6}$ | $10^{-30.3}$ | $10^{-58.2}$ | $10^{-27.0}$ |
| | 3 | $10^{-18.1}$ | $10^{-30.6}$ | $10^{-52.0}$ | $10^{-28.3}$ |

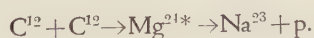
7×10^8 °K, and the reaction between C^{12} and O^{16} becomes effective at about $T = 10^9$ °K,

2) other reactions are not effective unless the temperature exceeds 10^9 °K.

In the above calculations we have assumed that the potential between two colliding nuclei is such that out of the channel radius there is only Coulomb repulsive potential and inside the channel radius a uniform nuclear potential. This assumption is not yet verified to be valid for reactions between heavy particles, so we want to discuss the validity of this assumption. For reactions between heavy nuclei the theory of fission may be useful, because our process can be inferred as an inverse process of fission^{*)}. To check our assumption, we will compare the penetration factor for the spontaneous fission of U^{230} obtained by the present method with the one obtained from the liquid-drop model of fission. The penetration factor e^{-G} is calculated as follows: $10^{-54.7}$ by the fission theory, $10^{-49.6}$ and $10^{-33.4}$ by the present method with $r_0 = 1.3$ and 1.4×10^{-13} cm, respectively. Then it is inferred that if r_0 is somewhat reduced, both theories lead to the same results. If the value of r_0 is replaced by 1.3×10^{-13} cm instead of 1.4×10^{-13} cm, we find that the reaction rate of $C^{12} + C^{12}$ with $E = 1$ Mev and $T = 10^9$ °K will increase by a factor about 4, namely the figure $10^{-8.4}$ year $^{-1}$ in Table III changes to $4 \times 10^{-8.4}$ year $^{-1}$.

We have also assumed that the angular momentum between colliding particles is zero. This assumption does not affect the results appreciably, because the above mentioned reaction rate $10^{-8.4}$ year $^{-1}$ for $l = 0$ is reduced only to $0.83 \times 10^{-8.4}$ year $^{-1}$, if we assume $l = 1$ state.

If the temperature becomes so high that the reactions among C^{12} , O^{16} , Ne^{20} , etc., themselves begin to contribute to the energy generation of stars, we observe that nuclei with mass number A which are not equal to $4n$ can be produced in these reactions, for example



These reactions may be important from the point of view of chemical abundance in the stars.

*) We thank Professor M. Taketani for his suggestion on this point.

§ 7. Discussions and further remarks

In this section we shall check the assumptions in the preceding sections, and discuss the results obtained.

The assumption that the momentum distributions of the reacting particles in the helium core are Maxwellian, is quite accurate, as the reactions proceed very slowly.

The first and the second assumptions in § 4 are regarded as reasonable from the present knowledge of stellar structure.

The third assumption that T_c , ρ_c and x_1 are constant, is not always valid during the successive stages of stellar evolution. As a star evolves, the values of these physical parameters will be altered according as the stellar structure changes. The results obtained in § 5 may be applicable to the stage of stellar evolution in which the structure does not appreciably change. The stars along the horizontal branch in the Hertzsprung-Russell diagram of the globular cluster might be identified as those at this stage. On the other hand, our results cannot be applied for stars whose structure changes considerably, e.g., the stars along the red giant branch in the globular cluster. In the latter case we must solve the equations (45) governing the change of chemical composition with reference to the stellar structure.

The results summarized in Figs. (10) ~ (19) show that the abundances of C^{12} , O^{16} and Ne^{20} in the hydrogen exhausted core depends critically on the exact reaction widths of these nuclei as discussed in § 5. But the general conclusion on the energy production in the core may be drawn from the closer examination of Figs. (20), and (21). As the reactions are considered to be occurring in the hydrogen exhausted core, the energy produced in the core decreases in proportion to X_α^3 if the reaction $3\alpha \rightarrow C^{12}$ alone contributes to the energy production. In the case of $X_\alpha > 0.5$, the energy produced in the core decreases approximately in proportion to X_α^3 , and this indicates that the reactions (II) and (III) do not contribute appreciably to the energy production. On the contrary, when $X_\alpha < 0.5$ the energy production is far greater than that given by the reaction (I) only, indicating that the reaction (II) and (III) play an important role in the energy production at this stage.

These results might directly be applied to the stars found in the horizontal branch of the Hertzsprung-Russell diagram of the globular cluster. The star of red giant branch of the globular cluster, however, increases its luminosity rapidly during its evolution. In the latter case, only a part of our results, viz., for $X_\alpha \simeq 1$, may be applicable. For a star whose luminosity is rapidly varying, we can proceed by the following procedure: Divide the time t or helium abundance X_α into small intervals and consider the stellar structure to remain unaltered during this interval. For each interval the variation of chemical composition and its effect on the luminosity can be calculated by the same method as in § 5. Then, proceed to the next interval of t or X_α using the stellar model constructed upon the various quantities calculated in the preceding interval. These rather laborious programs about the evolution of the giant stars are out of the scope of the present paper.

The results obtained in this paper give the qualitative suggestion about the general trend of nuclear reactions, in the later stage of stellar evolution.

Incidentally we will mention several desirable experiments for nuclear data which are necessary for the study of the nuclear reactions in stars. We want to know the exact location of the levels with about 7.65 Mev, 7.12 Mev and 4.95 Mev excitation energies of C^{12} , O^{16} , and Ne^{20} , respectively. Assignment of the spin value and the parity for the above level of Ne^{20} is also desirable. Further it would be very useful for measurement of the reaction widths: γ -width of the 7.65 Mev state of C^{12} , α -width and γ -width of the 7.12 Mev state of O^{16} and α -width of the 4.95 Mev state of Ne^{20} . For the further study of nuclear reactions in the star, detailed knowledges of the excited states of Mg^{24} , Si^{28} , etc., are also desirable.

We are deeply grateful to Professors M. Taketani, S. Nakamura and T. Hatanaka for their valuable advice and helpful discussions of many aspects of this problem. We are indebted to Miss. T. Yamada for her assistance in numerical calculations.

References

- 1) M. Taketani, T. Hatanaka and S. Obi, *Prog. Theor. Phys.* **15** (1956) 89.
- 2) E. J. Öpik, *Mem. Soc. Roy. Sci. Liege*, **13**, (1954), 131.
- 3) E. E. Salpeter, *Ap. J.* **115**, (1952), 326.
- 4) F. Hoyle, *Ap. J.*, Supplement II, No. 2, (1954).
- 5) F. Aizenberg and T. Lauritsen, *Rev. Mod. Phys.* **27**, (1955), 77.
- 6) D. W. Miller, V. K. Rasmussen and M. B. Sampson, *Phys. Rev.* **95** (1954), 649.
- 7) J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics*, (1952), 574.
- 8) S. Devons, *Proceedings of the 1954 Glasgow Conference*, p. 46.
- 9) T. Teichman and E. Wigner, *Phys. Rev.* **87**, (1952), 123.
- 10) M. Abramowitz, *J. Math. Phys.*, **33**, (1954), 111.
- 11) W. Shibagaki, 0.01% Tables of Modified Bessel Functions with the Account of the method used in the calculation, (1955).
- 12) L. Bloch, M. H. Hull, Jr. A. A. Broyles, W. G. Fourcicus, and B. Breit, *Rev. Mod. Phys.* **23** (1951), 147.
- 13) R. G. Freemantle, D. J. Prowse, A. Hossain and J. Rotbalt, *Phys. Rev.* **96**, (1954), 1270.
- 14) J. R. Holt: *Proceedings of the 1954 Glasgow Conference*, p. 62.
- 15) F. A. EL Bedewi and M. A. El Wahab, *Proc. Phys. Soc.* **68A**, (1955), 754.
- 16) A. R. Sandage and M. Schwarzschild, *Ap. J.*, **116**, (1952), 463.
- 17) F. Hoyle and M. Schwarzschild, *Ap. J. Supplement II*, (1955), no. 13.
- 18) L. H. Aller: *Astrophysics II*, (1954), Chap. 2.
- 19) *Mathematical Tables of the British Association for the Advancement of Science*, Vol. 2, (1932)
- 20) S. Obi: to be published in *Publ. Astr. Soc. Japan*, **8**, (1956).
- 21) P. Endt and J. Kluwyner, *Rev. Mod. Phys.* **26** (1954), 95.
- 22) S. Frankel and N. Metropolis, *Phys. Rev.* **72** (1947), 914

Letters to the Editor

A Lattice Model of Liquid Helium

Takeo Matsubara*

and

Hirotugu Matsuda**

*Research Institute for Fundamental Physics,
Kyoto University, Kyoto* and
Quantum Chemistry Laboratory, Chemistry
Department, Kyoto University, Kyoto***

July 25, 1956

Concerning the λ -transition of liquid helium II, the following three points seem to remain unexplained from the first principle:

- (a) The λ -transition is of the second kind.
- (b) The transition temperature T_λ decreases with increasing pressure p or density ρ , i.e.

$$dT_\lambda/dp < 0 \text{ or } dT_\lambda/d\rho < 0.$$

- (c) The specific volume under constant pressure increases with decreasing temperature below λ point.

We have worked out a lattice theory of liquid helium which enables us to solve these questions in a simple way.

Instead of considering continuum liquid, we assume that each atom in liquid helium can occupy one of lattice points, say simple cubic lattice points with a lattice constant d . Defining creation and annihilation operators of atom occupying i -th lattice point by a_i^* and a_i respectively, we approxi-

mate the total Hamiltonian of liquid helium as

$$\mathcal{H}_L = (\hbar^2/2md^2) \sum_{\langle ij \rangle} (a_i^* - a_j^*) (a_i - a_j) - v_0 \sum_{\langle ij \rangle} a_i^* a_i a_j^* a_j. \quad (1)$$

The first term of the right hand side of (1) represents the kinetic energy of the system and the second term the potential energy. $-v_0$ is the value of interatomic potential at the nearest neighbor distance. $\langle ij \rangle$ means to take summations over all nearest neighboring pair points. (1) can be derived from the true Hamiltonian of the system by passing from continuum to a discrete lattice liquid. In order to exclude the possibility that two atoms occupy a lattice point simultaneously, we impose on the operators a_i^* and a_i the restrictions

$$a_i^* a_i + a_i a_i^* = 1, \quad a_i a_i = a_i^* a_i^* = 0 \quad (2a)$$

in addition to the ordinary commutation relations for Bose particles:

$$[a_i, a_j] = [a_i^*, a_j^*] = [a_i, a_j^*] = 0 \quad (i \neq j). \quad (2b)$$

It is readily seen that the operators satisfying (2a) and (2b) are equivalent to a set of spin operators $\{S_i\}$, since it holds for $S=1/2$ that

$$S_{i+} S_{i-} + S_{i-} S_{i+} = 1, \quad S_{i-} S_{i-} = S_{i+} S_{i+} = 0, \quad (3a)$$

$$|S_{i-}, S_{j-}| = |S_{i+}, S_{j+}| = |S_{i-}, S_{j+}| = 0 \quad (i \neq j), \quad (3b)$$

where

$$S_{i\pm} = S_{iz} \pm iS_{iy}, \quad S_{iz} = S_{iz}, \quad S_{i-} = S_{iz} - 1/2. \quad (3c)$$

We may, therefore, expect that the system whose Hamiltonian is given by (1) will be mathematically equivalent to a certain

system of spins. Indeed, if we consider a ferromagnetic system with an anisotropic exchange coupling, subjected to an external magnetic field and let the Hamiltonian be

$$\mathcal{H}_F = -J \sum_{\langle ij \rangle} (S_{ix} S_{jx} + S_{iy} S_{jy}) - J' \sum_{\langle i \rangle} S_{iz} S_{jz} - H \sum_i S_{iz}, \quad (4)$$

then it is proved that the grand partition function of the lattice liquid

$$\Xi_L = \text{Trace} [\exp(-\alpha \sum_i a_i^* a_i - \beta \mathcal{H}_L)] \quad (5)$$

is equal to the partition function of the ferromagnets

$$\Xi_F = \text{Trace} [\exp(-\beta \mathcal{H}_F)], \quad (6)$$

except for a multiplication constant, that is,

$$\Xi_L = C \Xi_F, \quad (7)$$

provided that the following identification is made:¹⁾

$$\hbar^2/2md^2 = J/2, \quad v_0 = J',$$

$$\alpha = -\beta(H + z/2 J - z/2 J'),$$

$$C = \exp[N\beta(1/2 H - 1/8 z J')], \quad (8)$$

where N and z are the total number of lattice points and the number of nearest neighbor lattice points respectively. All other thermodynamical quantities of both systems can be easily identified with each other. It is easier and more familiar to deal with Ξ_F than with Ξ_L , because we have many well-established approximation methods to calculate Ξ_F —for instances, spin-wave approximation at very low temperatures and molecular field or Bethe approximation near and above the Curie temperature.

Having evaluated Ξ_F for the most important case of $J > J'$ through various approximations, and translating the results thus obtained in terms of lattice liquid, we

are led to many interesting conclusions. For instances:

(1) First of all, the λ -transition of lattice liquid corresponds to the ferromagnetic transition where a spontaneous magnetization appears in the xy plane, and it turns out to be of the second kind as should be expected. We have obtained specific heat of liquid which is proportional to T^3 at very low temperatures and accompanied by a jump at the λ -point. This is a correct answer to the question (a).

(2) The density (or pressure) dependence of the λ -point is the same as that stated in (b) if we assume that there are more atoms than holes in our model. This is due to the fact that the stronger the applied magnetic field, the lower is the Curie temperature. It may correspond to the fact that the more holes exist in the lattice, the easier the superfluidity can set in.

The details of the content of this letter will be published shortly.

- 1) A similar correspondence existing between lattice gas and Ising model of ferromagnets was pointed out by Yang and Lee. But, since they did not take into consideration the vector character of spins, they failed to get the quantum effect of the lattice gas. See Phys. Rev. 87 (1952), 410.

On the Effects of Heavy Unstable Particles to the Anomalous Magnetic Moments of Nucleons

Chikashi Iso* and Syurei Iwao**

*Department of Physics, Tokyo University of Education, Tokyo, and

**Department of Physics, Osaka City University, Osaka

June 21, 1956

The anomalous magnetic moments of nucleons have been calculated by many authors, but it is found very difficult to fit the calculated values to the experimental ones

$$\Delta\mu_p = 1.790, \quad \Delta\mu_n = -1.910,$$

$$\Delta\mu_p + \Delta\mu_n = -0.120$$

by means of the ordinary π -meson theories. Above all $|\Delta\mu_p + \Delta\mu_n|$ calculated is very much larger than the observed value.

There is a very reason to believe that such a situation as mentioned above comes from the following:

1) In the direct coupling π -meson theory, it is found that $|\Delta\mu_p + \Delta\mu_n|$ is given only by the contribution from the nucleon current, in which higher momentum parts have considerable effects. This is caused by the pair effect due to the γ_5 matrix. Cutting off the contribution from higher momentum gives rather reasonable values, but $|\Delta\mu_p + \Delta\mu_n|$ is still too large.¹⁾

(See Table 1.)

2) We know that there exist heavy mesons (θ , τ , etc.) which also strongly interact with nucleons, so we have to take into account the effects of these mesons. According to the assignment of the isotopic spins of heavy mesons and hyperons,²⁾ the contribution to $|\Delta\mu_p + \Delta\mu_n|$ from hyperon current vanishes and only that from heavy mesons remains. This may probably leads to the more reasonable value of $|\Delta\mu_p + \Delta\mu_n|$.

After the previous work³⁾ masses, spins and effective coupling constants of heavy unstable particles have gradually become clear, so it is desirable to give the order of the calculated values of the anomalous magnetic moments of nucleons due to the heavy mesons.

In the following, it is assumed that the spins of hyperons and heavy mesons are $1/2$ and 0 respectively and that they have direct coupling with nucleons. The calculations have been done in the second order of coupling constant among hyperons, heavy mesons and nucleons, and the results are expressed in unit of the nuclear magneton. We assume the isotopic spin of heavy mesons I_K to be $1/2$ and that of hyperons I_H to be 0 (In this case the coupling constants are taken as $f_0 \doteq 0$, $f_1 = 0$) or 1 ($f_0 = 0$, $f_1 \doteq 0$). In the above brackets suffices of the coupling constants represent the isotopic spins of hyperons. When $I_K = 0$ or 1 heavy mesons and $I_H = 1/2$ hyperons are found, B_1 and B_2 in the following equations must be exchanged.

i) Even parity for hyperon and even parity for heavy meson:

$$\Delta\mu_p = \frac{2f_1^2}{4\pi} \frac{B_1}{2\pi} + \frac{f_1^2 + f_0^2}{4\pi} \frac{B_2}{2\pi},$$

$$\Delta\mu_n = -\frac{2f_1^2}{4\pi} \frac{B_1}{2\pi} + \frac{2f_1^2}{4\pi} \frac{B_2}{2\pi}$$

where B_1 and B_2 are caused by hyperon current and meson current respectively and are given by

$$\begin{aligned} B_1 &= (\kappa^2 + \kappa - \lambda^2 - 1/2) + 1/2 \times \\ &\times \{ \lambda^2 - (\kappa^2 - 1 - \lambda^2)(1 + \kappa) - (\kappa^2 - 1 - \lambda^2)^2 \} \\ &\quad \times \log(\kappa^2/\lambda^2) \\ &- [1/2 \{ \lambda^2 - (\kappa^2 - 1 - \lambda^2)(1 + \kappa) \\ &\quad - (\kappa^2 - 1 - \lambda^2)^2 \} (\kappa^2 - 1 - \lambda^2) \\ &\quad + \lambda^2(\kappa^2 + \kappa - \lambda^2)] \\ &\times \frac{2}{\sqrt{4\lambda^2 - (\kappa^2 - 1 - \lambda^2)^2}} \times \cos^{-1} \frac{\kappa^2 + \lambda^2 - 1}{2\kappa\lambda}, \\ B_2 &= 1/2 - (\lambda^2 - \kappa^2 - \kappa) \end{aligned}$$

$$\begin{aligned}
& -1/2 \{ (\lambda^2 - \kappa^2 - 1) (\lambda^2 - \kappa^2 - \kappa) \\
& \quad - (\kappa^2 - \kappa) \} \log \kappa^2 / \lambda^2 \\
& - [1/2 \{ (\lambda^2 - \kappa^2 - 1) (\lambda^2 - \kappa^2 - \kappa) - \\
& (\kappa^2 + \kappa) \} (\lambda^2 - \kappa^2 - 1) + \kappa^2 (\kappa^2 + \kappa - \lambda^2)] \\
& \times \frac{2}{\sqrt{4\kappa^2 - (\lambda^2 - 1 - \kappa^2)^2}} \cos^{-1} \frac{\kappa^2 + \lambda^2 - 1}{2\kappa\lambda}
\end{aligned}$$

in which

$\kappa = (\text{hyperon mass}) / (\text{nucleon mass})$.

$\lambda = (\text{heavy meson mass}) / (\text{nucleon mass})$.

ii) Even parity for hyperon and odd parity for heavy meson :

The results can be obtained from case i) by changing the sign of κ and B.

iii) Odd parity for hyperon and even parity for heavy meson :

The results are the same as the case ii).

iv) Odd parity for hyperon and odd parity for heavy meson :

The results are the same as the case i).

Using the observed values of masses, $m_\pi = 273m_e$, $m_K = 966m_e$, $m_\Lambda = 2182m_e$, $m_\Sigma = 2327m_e$ and the coupling constants⁴⁾

$$f^2(N\Lambda K)/4\pi \sim 1, \quad f^2(N\Sigma K)/4\pi \sim 0.4,$$

$$f^2(NN\pi)/4\pi \sim 15,$$

the results of the calculation are listed in Table 1.

Table 1. In this table θ meson is assumed to be scalar and τ meson to be pseudoscalar.

| Spins and parities of Hyperons | 1/2, + | 1/2, - |
|---|--|--|
| $\theta\Lambda$ | $\Delta\mu_p = -0.086 f^2/4\pi \sim -0.09$ $\Delta\mu_n = 0 \quad \quad \quad = 0$ | $\Delta\mu_p = 0.029 f^2/4\pi \sim 0.03$ $\Delta\mu_n = 0 \quad \quad \quad = 0$ |
| $\theta\Sigma$ | $\Delta\mu_n = 0.090 f^2/4\pi \sim 0.04$ $\Delta\mu_n = -0.327 f^2/4\pi \sim -0.13$ | $\Delta\mu_p = -0.359 f^2/4\pi \sim -0.14$ $\Delta\mu_n = 0.445 f^2/4\pi \sim 0.18$ |
| $\tau\Lambda$ | The values are obtained by changing the parity of hyperon Λ in the case of $\theta\Lambda$. | |
| $\tau\Sigma$ | The values are obtained by changing the parity of hyperon Σ in the case of $\theta\Sigma$. | |
| πN (lowest order, renormalization) | $\Delta\mu_p = 0.035 f^2/4\pi \sim 0.53, \quad \Delta\mu_n = -0.262 f^2/4\pi \sim -3.93$ | |
| πN (cut off at nucleon mass) | $\Delta\mu_p = 0.041 f^2/4\pi \sim 0.61, \quad \Delta\mu_n = -0.013 f^2/4\pi \sim -1.9$ | |

Now, one might think that there should be considerable contributions from higher momentum parts and that the cut-off procedure would suppress the contributions from heavy unstable particles, because of the heavy masses of heavy mesons and hyperons. But the cut-off procedure mainly affects the hyperon current, and $|\Delta\mu_p + \Delta\mu_n|$ caused by heavy mesons will not be

suppressed so much.

The contributions of heavy unstable particles to the anomalous magnetic moment of nucleons are 10~20%. and these effects can not be neglected for the consideration of $|\Delta\mu_p + \Delta\mu_n|$. It must be remarked that the relatively small contributions from the heavy mesons are ascribed to the smallness of the effective coupling constants rather

than the largeness of their masses. (The smallness of the coupling constants and the largeness of their masses may be connected more closely in future theory.)⁵⁾

The authors wish to express their sincere thanks to Mr. M. Sato for his careful reading of the manuscript.

- 1) S. Goto, Prog. Theor. Phys. **12** (1954), 699; H. Hasegawa, S. Matsuyama and T. Akiba, Prog. Theor. Phys. **12** (1954), 548; N. Fukuda, Soryushiron Kenkyu (mimeographed circular in

Japanese) **7** (1954), 142.

- 2) T. Nakano and K. Nishijima, Prog. Theor. Phys. **10**(1953), 581.
 - 3) C. Iso, Soryushiron Kenkyu, **6** (1954), 1040.
 - 4) S. Iwao, Prog. Theor. Phys. **13**(1955), 111 and Soryushiron Kenkyu. **10**(1955), 192; S. Minami, to be published: The authors would like to thank Dr. S. Minami for informing us on the results of calculations for the $K^+ - N$ scattering in advance of publication.
 - 5) As suggested by the thermodynamical treatment of the high energy phenomena; see for example, S. Z. Belenkij and L. D. Landau, Usp. Fiz. Nauk USSR. **56** (1955), 309.
-

On a Solution of Field Equations in Einstein's Unified Field Theory, I*

N. N. GHOSH

Department of Pure Physics, Calcutta University, Calcutta, India

(Received February 13, 1956)

This paper is concerned with the general solution of field equations (in the 'strong' form) corresponding to a non-symmetric tensor field $g_{\mu\nu}$ whose symmetric components s and anti-symmetric components a are expressed by using general indices k, m, n, l instead of 1, 2, 3, 4. The components s_{kk}, s_{ll}, a_{kl} are taken to be functions of x_k while s_{mm}, s_{nn}, a_{mn} are each expressible as product of a function of x_k and x_m . In the process of solution of the relevant field equations there appears a parameter λ in an auxiliary equation which can take negative, positive or zero values. It is found that the usual spherically symmetric solutions considered by Eonnor result when λ is negative. The other two types of solution are also noted in this connection and a particular case not considered by Bonnor in his paper is also discussed.

§ 1. Introduction

In a recent paper (Ghosh¹⁾) solution of the 64 I 's defined by the system of equations

$$g_{\mu\nu,\sigma} = g_{\mu\alpha} \Gamma_{\sigma\nu}^{\alpha} + g_{\alpha\nu} \Gamma_{\mu\sigma}^{\alpha}$$

together with the computed values of the 16 components of the Hermitianized Einstein tensor $R_{\mu\nu}$ was obtained corresponding to a non-symmetric tensor field $g_{\mu\nu}$ whose symmetric components s and anti-symmetric components a are represented by $s_{kk}, s_{mm}, s_{nn}, s_{ll}, a_{kl}, a_{mn}, a_{nm}, a_{lk}$, where k, m, n, l denote distinct indices 1, 2, 3, 4 taken in different orders. Adopting these general indices with no summation convention involved in them the purpose of the present paper is to solve the field equations in the so-called 'strong' form, that is, of the system of equations

$$R_{\mu\nu} = 0,$$

$$R_{\mu\nu} = 0, \quad (\mu, \nu = k, m, n, l)$$

$$\Gamma_{\mu} = \Gamma_{\mu\nu}^{\nu} = 0$$

for the case in which

s_{kk}, s_{ll}, a_{kl} are functions of x_k only and

s_{mm}, s_{nn}, a_{mn} are functions of x_k and x_m of the type

*) An abstract of this paper was read at the 43rd session of the Indian Science Congress, 1956.

$$\begin{aligned}
 s_{mm} &= \phi_{mn}(x_k) \phi'_{mn}(x_m), \\
 s_{nn} &= \phi_{nn}(x_k) \phi'_{nn}(x_m), \\
 a_{mn} &= \phi_{mn}(x_k) \phi'_{mn}(x_m).
 \end{aligned}
 \tag{1.2}$$

A special case of this type has been considered by Bonnor³⁾, in connection with spherically symmetric field for which $(k, m, n, l) = (1, 2, 3, 4)$ with polar coordinates $x_k = r$, $x_m = \theta$, $x_n = \phi$, $x_l = t$. The present treatment is more general than Bonnor's and throws light on some points. In § 3 we have completed the solution of a particular case noticed by Bonnor but not treated by him as it is not compatible with spherically symmetric solutions. However, the solution has been proved to be trivial. In § 4 it is found that spherical symmetry in the field is involved in Einstein's fundamental equations (1.1) in consequence of a parameter λ in the derived equation (4.7) or (4.8) admitting negative values. The types of solution for positive values of λ and for $\lambda = 0$ have also been noted in this connection and general solutions for these cases have been given. As we have retained the general indices k, m, n, l throughout, these general solutions also apply to other two types of tensor field noted in my previous paper.

§ 2. Expressions for the Γ and simplification of field equations

Our fundamental tensor $g_{\mu\nu}$ satisfies the conditions (15) of my previous paper¹⁾ by virtue of (1.2), therefore we can make use of the formulae (25) of (I) for calculating Γ 's. We note that out of the 12 quantities A 's, B 's, H 's only 4 survive and we write them in the following abbreviated form

$$\begin{aligned}
 A_k &= (\gamma_{mkn} - c_m^n c_n^m \gamma_{mnk}) / (1 - c_m^n c_n^m), \\
 B_k &= (\gamma_{mkn} - \gamma_{mnk}) / (1 - c_m^n c_n^m), \\
 H_k &= (\gamma_{kkk} + \gamma_{llk} - 2\gamma_{klk}) / (1 - c_k^l c_l^k), \\
 H_m &= (\gamma_{mmm} + \gamma_{llm} - 2\gamma_{mlm}) / (1 - c_m^l c_l^m).
 \end{aligned}
 \tag{2.1}$$

In the above, symbols like $\gamma_{\mu\lambda k}$, $\gamma_{\mu\nu m}$ denote respectively $s_{\mu\lambda, k}/s_{\mu\lambda}$ and $a_{\mu\nu, m}/a_{\mu\nu}$ while symbols like c_ν^μ stand for $a_{\mu\nu}/s_{\mu\lambda}$.

The non-vanishing Γ 's are then written in the following 9 groups of two each:

$$\begin{aligned}
 \Gamma_{kn}^m &= -(1/2) c_n^m B_k, & \Gamma_{ln}^m &= (1/2) c_n^m c_l^k B_k, \\
 \Gamma_{lm}^m &= -(1/2) c_l^k A_k, & \Gamma_{km}^m &= (1/2) A_k, \\
 \Gamma_{mm}^k &= -\frac{1}{2} \frac{s_{mm}}{s_{kk}} (A_k + c_m^n c_n^m B_k), \\
 \Gamma_{kl}^k &= -(1/2) c_l^k H_k, & \Gamma_{kl}^l &= (1/2) \gamma_{llk} + (1/2) c_k^l c_l^k H_k, \\
 \Gamma_{kk}^k &= (1/2) \gamma_{kkk}, & \Gamma_{ll}^k &= -\frac{s_{ll}}{s_{kk}} ((1/2) \gamma_{llk} + c_k^l c_l^k H_k),
 \end{aligned}
 \tag{2.3}$$

with a single one

$$\Gamma_{mn}^k = - (1/2) \frac{a_{mn}}{s_{kk}} (A_k + B_k).$$

It may be noted that the unwritten member in each of the first five follows by an interchange of indices $m \leftrightarrow n$ while in the next four by an interchange $k \leftrightarrow m$, $l \leftrightarrow n$.

Further we have

$$\begin{aligned} 1 - c_k^l c_l^k &\neq 0, & 1 - c_m^n c_n^m &\neq 0, \\ (1 - c_k^l c_l^k)^2 + 4 c_k^l c_l^k (1 - c_m^n c_n^m) &\neq 0, & (2.4) \\ (1 - c_m^n c_n^m)^2 + 4 c_m^n c_n^m (1 - c_k^l c_l^k) &\neq 0, \end{aligned}$$

so that the above values of the Γ are unique. The relevant field equations in (1.1) can all be constructed from the simplified expressions for $R_{\mu\nu}$ and $R_{\mu\nu}$ given in my previous paper. These are

$$R_{kk}=0, R_{ll}=0, R_{mn} = \frac{s_{mn}}{s_{nn}} R_{nn}=0, R_{kl}=0, R_{mn}=0. \quad (2.5)$$

On account of the last equation in (1.1) we have the conditions $\Gamma_l=0$, $\Gamma_n=0$ yielding the relations

$$H_k = 2A_k, \quad H_m = 0. \quad (2.6)$$

Referring to (2.1) we can easily see that the condition $H_m=0$ is satisfied by the field components in (1.2) if $\phi_{mn} = \sqrt{(\phi_{mn} \phi_{nn})}$. Substituting from (1.2), A_k, B_k in (2.1) are now expressible as

$$A_k = \frac{\phi_{mm} \phi_{mn,k} + \phi_{mn} \phi_{mn,k}}{\phi_{mn}^2 + \phi_{mn}^2}, \quad (2.7)$$

$$\frac{\phi_{mn}}{\phi_{mn}} B_k = \frac{\phi_{mn} \phi_{mn,k} - \phi_{mn} \phi_{mn,k}}{\phi_{mn}^2 + \phi_{mn}^2}. \quad (2.8)$$

Let us now proceed to the simplification of the system of equations (2.5) by proper substitutions of the Γ 's from (2.3) making use of the relations (2.6).

Utilising the relation $H_k = 2A_k$ that is,

$$\Gamma_{kk}^k + A_k + \Gamma_{kl}^l = \gamma_{kkk} - \gamma_{kkk}$$

the equation $R_{ll}=0$ reduces to

$$-\frac{\partial}{\partial x_k} \left(\frac{s_{kk}}{a_{kl}} \Gamma_{ll}^k \right) + \frac{1}{2} \frac{a_{kl}}{s_{kk}} (3A_k^2 - c_m^n c_n^m B_k^2) = 0. \quad (2.9)$$

The equation $R_{kl}=0$ gives

$$\frac{\partial}{\partial x_k} \left(-\frac{a_{kl}}{s_{kk}} A_k \right) - \frac{1}{2} \frac{a_{kl}}{s_{kk}} (3A_k^2 - c_m^n c_n^m B_k^2) = 0. \quad (2.10)$$

Adding the above and simplifying we get

$$\frac{\partial}{\partial x_k} \left(\frac{s_{ll}}{a_{kl}} \Gamma_{kl}^l \right) = 0 \quad (2 \cdot 11)$$

as obtained by Bonnor.

Combining now the equations $R_{kk}=0$, $R_{ll}=0$ in

$$R_{kk} - \frac{s_{kk}}{s_{ll}} R_{ll} = 0 \quad (2 \cdot 12)$$

we obtain by a straight-forward calculation

$$(1 - c_k' c_l^k) (-A_{k,k} + (1/2) A_k^2 + (1/2) c_m^n c_n^m B_k^2 + A_k \gamma_{klk}) = 0. \quad (2 \cdot 13)$$

Eliminating $A_{k,k}$ between the equations (2·10) and (2·13) we readily arrive at the equation*

$$A_k \Gamma_{kl}^l = 0. \quad (2 \cdot 14)$$

The two remaining equations $R_{mm}=0$, $R_{nn}=0$ yield respectively

$$\begin{aligned} & A_{k,k} + c_m^n c_n^m B_{k,k} + (A_k + c_m^n c_n^m B_k) (\gamma_{mmk} - (1/2) \gamma_{kkk} + \Gamma_{kl}^l) \\ & + c_m^n c_n^m B_k (A_k - B_k + 2c_m^n c_n^m B_k) = - \frac{2s_{kk}}{s_{mm}} f(x_m), \end{aligned} \quad (2 \cdot 15)$$

$$\begin{aligned} & A_{k,k} + B_{k,k} + (A_k + B_k) (\gamma_{nnk} - (1/2) \gamma_{kkk} + \Gamma_{kl}^l) \\ & + B_k (A_k + c_m^n c_n^m B_k) = 0, \end{aligned} \quad (2 \cdot 16)$$

where $f(x_m)$ stands for the expression

$$(1/2) [(\gamma_{nnm})_{,m} - (1/2) \gamma_{nnn} (\gamma_{mmm} - \gamma_{nnm})]. \quad (2 \cdot 17)$$

§ 3. Solution of field equations with $A_k=0$

From (2·14) it follows that there are two cases to consider according as (i) $A_k=0$ (ii) $\Gamma_{kl}^l=0$.

Proceeding with case (i)** and referring to (2·13) we have also $B_k=0$. Hence it follows from (2·1) that s_{mm} , s_{nn} , a_{nn} must be functions of x_m only. The equation (2·11) gives now

$$\partial / \partial x_k (s_{ll,k} / a_{kl}) = 0. \quad (3 \cdot 1)$$

Further we have the relation $H_k=0$, that is,

$$\partial / \partial x_k (a_{kl} / \sqrt{s_{kk} s_{ll}}) = 0. \quad (3 \cdot 2)$$

*) For an alternative way of obtaining this see Bonnor²⁾.

**) This case was not considered by Bonnor as being incompatible with spherically symmetric solutions.

Integrating the above equations we get

$$\begin{aligned}a_{kl} &= c_1 \phi'(x_k) e^{\phi(x_k)}, \\s_{kl} &= c_2 \{\phi'(x_k)\}^2 e^{\phi(x_k)}, \\s_{ll} &= e^{\phi(x_k)},\end{aligned}\tag{3.3}$$

where $\phi(x_k)$ is an arbitrary function of x_k and c 's are arbitrary constants.

Again putting $A_k=0$, $B_k=0$ in (2.15) we have $f(x_m)=0$ which in conjunction with the relation $H_m=0$ yields

$$\partial/\partial x_m (s_{mn,m}/a_{mn})=0,\tag{3.4}$$

$$\partial/\partial x_m (a_{mn}/\sqrt{s_{mm}s_{nn}})=0.\tag{3.5}$$

Thus the components a_{mn} , s_{mm} , s_{nn} satisfy equations of the same type as (3.1) and (3.2) with indices k, l replaced by m, n respectively. The solution thus obtained is, however, trivial for it appears that the set of components a_{kl} , s_{kl} , s_{ll} is transformed into constants by applying the coordinate transformation

$$\begin{aligned}e^{1/2 \cdot \phi(x_k)} \cos(\omega x_l) &= x'_l, \\e^{1/2 \cdot \phi(x_k)} \sin(\omega x_l) &= x'_l,\end{aligned}\tag{3.6}$$

the constant ω being suitably chosen. Similar remark applies to the other set of components a_{mn} , s_{mm} , s_{nn} which also reduce to constants.

§ 4. Solution of field equations with $\Gamma_{kl}^i=0$

Let us now consider case (ii). Since $\Gamma_{kl}^i=0$ we have from (2.3)

$$\frac{1}{2} \gamma_{lk} = -c_k^l c_l^k A_k.\tag{4.1}$$

The relations (2.1) and (2.6) give therefore

$$A_k = \frac{1}{2} \gamma_{kk} - \gamma_{klk}.\tag{4.2}$$

From (2.7) we have also

$$A_k = \frac{1}{2} \frac{\partial}{\partial x_k} \log(\phi_{mm}^2 + \phi_{nn}^2).\tag{4.3}$$

Integrating (4.2) and (4.3) we get

$$\phi_{mm}^2 + \phi_{nn}^2 = \frac{s_{kk}}{a_{kl}^2} \cdot \text{const}.\tag{4.4}$$

Next, we take up equations (2.15) and (2.16). Eliminating $A_{k,k}$ and $B_{k,k}$ in succession between these equations we find

$$(1 - c_m^n c_n^m) [B_{k,k} - B_k^2 (1 - c_m^n c_n^m) + B_k (A_k - \frac{1}{2} \gamma_{kk})] = (2s_{kk}/s_{nm}) f(x_m),\tag{4.5}$$

$$(1 - \epsilon_n^m \epsilon_n^m) [A_{k,k} + A_k (A_k - \frac{1}{2} \gamma_{kkk})] = - (2s_{kk}/s_{mm}) f(x_m). \quad (4.6)$$

Observing now that the left-hand side in each of the above is a function of x_k only while in the right-hand side there is a factor $2f(x_m)/\zeta_{mm}^h$ which is a function of x_m only, we infer that, for consistency, we must have the following derived equation

$$2f(x_m) = \lambda/2 \cdot \zeta_{mm}^h,$$

where λ is some constant. Simplifying $f(x_m)$, given in (2.17), we write the above as

$$\frac{\lambda}{2} \zeta_{mm}^h = \frac{\zeta_{nn,mm}^h}{\zeta_{nn}^h} - \frac{1}{2} \frac{\zeta_{nn,m}^h}{\zeta_{nn}^h} \left(\frac{\zeta_{mm,m}^h}{\zeta_{mm}^h} + \frac{\zeta_{nn,m}^h}{\zeta_{nn}^h} \right), \quad (4.7)$$

an equation involving ζ_{mm}^h and ζ_{nn}^h .

Effecting a separation of variables, after integrating once, we can rewrite (4.7) in the form

$$\zeta_{mm}^{1/2} = \frac{2(\zeta_{nn}^{1/2})_{,m}}{(\lambda \zeta_{nn}^h + \mu)^{1/2}}, \quad (4.8)$$

where μ is a constant of integration which we assume to be positive. This equation suggests that by a suitable coordinate transformation s_{mm} goes over to a function ϕ_{mm} of x_k only while ζ_{nn}^h in s_{nn} takes different forms according to the three cases (i) λ negative, (ii) λ positive, (iii) λ zero. We give below the explicit forms which (4.8) takes in different cases :

$$\begin{aligned} \text{(i)} \quad & \sqrt{-\lambda}/2 \cdot \int \zeta_{mm}^{1/2} dx_m = \sin^{-1} \left(\frac{\sqrt{-\lambda}}{\sqrt{\mu}} \zeta_{nn}^{1/2} \right), \\ \text{(ii)} \quad & \sqrt{\lambda}/2 \cdot \int \zeta_{mm}^{1/2} dx_m = \sinh^{-1} \left(\frac{\sqrt{\lambda}}{\sqrt{\mu}} \zeta_{nn}^{1/2} \right), \\ \text{(iii)} \quad & \sqrt{\mu}/2 \cdot \int \zeta_{mm}^{1/2} dx_m = \zeta_{nn}^{1/2}. \end{aligned} \quad (4.9)$$

The first is associated with the usual spherically symmetric field. The second gives a representation in which $\sin x_m$ in the first case is replaced by $\sinh x_m^*$. The third is the same as that discussed in (3.4), (3.5) since $f(x_m) = 0$, and gives ultimately constant values for ζ_{mm}^h and ζ_{nn}^h . The representation of ζ_{nn}^h in this case is one in which all the components are functions of x_k only.

Let us now proceed with the case $\lambda \neq 0$. Introducing $B_k = (\phi_{mm}/\phi_{mm}) B_k$ given in (2.8) we reduce the equation (4.5) to the form

$$\bar{B}_{k,k} + \bar{B}_k (A_k - (1/2) \gamma_{kkk}) = - \frac{s_{kk} \phi_{mm}}{\phi_{mm}^2 + \phi_{nn}^2} \cdot \frac{\lambda}{2} \quad (4.10)$$

while (4.6) gives

$$A_{k,k} + A_k (A_k - (1/2) \gamma_{kkk}) = - \frac{s_{kk} \phi_{mm}}{\phi_{mm}^2 + \phi_{nn}^2} \cdot \frac{\lambda}{2}. \quad (4.11)$$

* The symmetry characterized by this change may be called "Pseudo-spherical".

By virtue of the relations (4.2) and (4.4) we further reduce the above to

$$\bar{B}_{k,k} - \bar{B}_k \gamma_{klk} = 2c\phi_{mm} a_{kl}^2, \quad (4.12)$$

$$A_{k,k} - A_k \gamma_{klk} = -2c\phi_{mm} a_{kl}^2, \quad (4.13)$$

where c is an arbitrary constant in which λ is absorbed. Adopting the method of Wyman³ to solve the above equations in conjunction with (2.13) the final result may be presented in the form

$$\begin{aligned} \phi_{mm} + i\phi_{mm} &= h \operatorname{sech}^2(h^{1/2}x + a) / (-ic), \\ dx/dx_k &= a_{kl}, \end{aligned} \quad (4.14)$$

where a is an arbitrary complex constant, h is a pure imaginary constant and x an arbitrary function of x_k . Returning now to the relations (4.1), (4.4) along with (4.9) we get

$$\begin{aligned} s_{kk} &= c_1 (dx/dx_k)^2 (\phi_{mm}^2 + \phi_{mn}^2), \\ s_{ll} &= -1/c_1 \cdot (\phi_{mm}^2 + \phi_{mn}^2)^{-1} + c_2, \end{aligned} \quad (4.15)$$

$$\begin{aligned} \text{with} \quad \phi_{mm} &= 1, \quad \phi_{nn} = \sin^2 x_m \quad (\text{if } \lambda \text{ negative}) \\ &= \sinh^2 x_m \quad (\text{if } \lambda \text{ positive}), \end{aligned}$$

where the c 's are arbitrary real constants. The set of solutions (4.14) and (4.15) with arbitrary function and constants adjusted according to the natural (non-covariant) boundary condition gives Bonnor's solution. Recently its covariant form has been obtained by Ikeda.⁴

When $\lambda=0$ the components of the field tensor $g_{\mu\nu}$ are all functions of x_k . Referring to (4.12) and (4.13) we have now the equations

$$\bar{B}_{k,k} - \bar{B}_k \gamma_{klk} = 0, \quad (4.16)$$

$$A_{k,k} - A_k \gamma_{klk} = 0. \quad (4.17)$$

Using (4.17) in (2.13) we have also

$$A_k - \bar{B}_k = 0.$$

Hence we get on integration

$$A_k = \pm \bar{B}_k = \alpha a_{kl},$$

where α is a constant of integration.

Making use of the relation

$$d/dx_k \cdot \log(\phi_{mm} + i\phi_{mm}) = A_k + i\bar{B}_k, \quad (4.19)$$

which follows from (2.7) and (2.8) we easily obtain the general solution as

$$\begin{aligned} a_{kl} &= dx/dx_k, \\ s_{mm} &= s_{nn} = [h_0 \sin(\alpha x) + h_1 \cos(\alpha x)] e^{\pm \alpha x}, \\ a_{mm} &= [h_0 \cos(\alpha x) - h_1 \sin(\alpha x)] e^{\pm \alpha x}, \end{aligned} \quad (4.20)$$

where h 's and α are arbitrary constants, x being an arbitrary function of x_k .

The solutions for s_{kk} , s_{ll} are given by (4.15). These are now

$$\begin{aligned} s_{kk} &= c_1 (dx/dx_k)^2 (s_{mm}^2 + a_{mn}^2), \\ s_{ll} &= -(1/c_1) \cdot (s_{mm}^2 + a_{mn}^2)^{-1} + c_2. \end{aligned} \quad (4.21)$$

The author wishes to thank Prof. S. N. Bose for his kind interest and advice. Thanks are also due to the Ministry of Education, Government of India, for the award of a research scholarship.

References

- 1) N. N. Ghosh, *Prog. Theor. Phys.* **13** (1955), 587
- 2) W. Ponnor, *Proc. Roy. Soc. A* **210** (1952), 427.
- 3) M. Wyman, *Canad. J. Math.* **2**(1950), 427.
- 4) M. Ikeda, *Prog. Theor. Phys.* **15** (1956), 5.

Deuteron Stripping and the Compound Nucleus

Masayuki NAGASAKI

Department of Physics, Tokyo University of Education, Tokyo

(Received May 24, 1956)

Deuteron stripping (d, p) reactions are investigated as corrections to the ordinary formal theories of nuclear resonance reactions, based on the compound nucleus model, in which the decaying tails of the internal wave functions of various reaction products are neglected. The amplitude for a (d, p) stripping reaction is shown to be due to the overlapping of the decaying tail of the deuteron internal wave function and the wave function of the neutron trapped in the target nucleus. It differs from the Butler's stripping amplitude in that the contributions from the configurations in which the proton is within the range of the interaction between the proton and the target nucleus or between the proton and the neutron interacting with the target nucleus, are properly included in the amplitude for the (d, p) reaction due to the formation and the subsequent decay of the compound nucleus. The effects of the compound nucleus formation on the angular distribution of emerging protons are also discussed in connection with the interpretations of some phenomenological calculations hitherto presented. As the formulation is given in a general way, it can be applied also to other cases than the deuteron stripping reaction, such as (p, d) pick-up, (n, p) direct or (t, d) stripping reactions.

§ 1. Introduction

Since Butler¹⁾ proposed an approximate approach to (d, p) and (d, n) reactions aiming at the explanation of the experimental observations of peaked forward angular distributions of emerging nucleons, numerous theoretical studies¹⁾⁻¹¹⁾ have been performed on the theory of deuteron stripping or pick-up reactions. In many of these theories, the main matter of concern has been a simpler derivation of the angular distribution of stripped-off nucleons from the Born approximation neglecting the interaction between the target nucleus and the nucleon to be stripped-off, as compared with the complicated procedure of boundary condition matching used by Butler. Or it has been the calculation of the effects of the final state interaction between the stripped-off nucleons and the residual nucleus on the angular distribution. As to the effects of the compound nucleus formation, however, only a few investigations have been hitherto done.¹²⁾

In view of the great successes of the compound nucleus model in explaining the main features of nuclear reactions at low and intermediate energy regions, it seems important to investigate how the deuteron stripping reactions can be related to the deuteron induced reactions which proceed through the compound nucleus formation. This investigation is important also because the compound nucleus contributions to these deuteron reactions are expected to modify greatly the angular distribution of emerging nucleons in some cases, and therefore the values of the reduced width of the captured nucleon in the

residual nucleus which are obtained by fitting the theoretical angular distribution to experimental observations. The purpose of this note is to give, though formally, a unified expression to the reaction amplitudes which involves the contributions both from the compound nucleus formation and the direct stripping process, by considering the latter contributions as corrections to the ordinary formal theories of nuclear resonance reactions.

As is well known, there have been the formal theories of nuclear resonance reactions of Kapur and Peierls¹³⁾ and of Wigner and his collaborators¹⁴⁾ among those based on the compound nucleus model.^{*)} In these formalisms the configuration space of all the constituent nucleons of the system are divided into two parts. One is the internal region, and the other is the external region. There are two alternative ways of defining the internal and the external regions. For instance, we may define the internal region as the part of the configuration space in which the internal wave functions of various pairs of two particles, into which the system decays, overlap with one another, and the external region as the residual part of the configuration space. (See Fig. 1.) If we choose this definition of the internal and external regions, these formalisms could be applied to the deuteron stripping reactions without any correction.¹⁵⁾

However, if some of the internal wave functions of pairs extend widely beyond the ranges of their internal interactions, the internal region so defined has no close relationship to the notion of the compound nucleus. In such a case, the internal region includes some parts of the configuration space in which all the strong interactions among the constituent nucleons of the system do not always act effectively to form compound states. For example, at the configuration in which the distance between the centers of mass of the deuteron and the target nucleus is somewhat smaller than the sum of a half the extension of the deuteron internal wave function plus the radius of the target nucleus (the shaded region in Fig. 1), the interactions between the proton and target nucleus, the proton and neutron, and the neutron and target nucleus do not always act with one another, though this configuration belongs to the internal region. Thus with this definition of the internal region, we could hardly say that in the internal region the energy of the incident particle is rapidly shared among all the constituent nucleons of the system. The internal region defined in this way therefore does not correspond to the compound nucleus.

Alternatively we may define the internal region as the part of the configuration space where all the strong interactions among the constituent nucleons of the system do act effectively to form a chaotic state, in which we could hardly trace out such modes of motion that resemble free motions of various pairs of two particles. In what follows, we shall choose this definition of the internal region, since this choice of the internal region has close relationship to the compound nucleus. The internal region so defined is identical to the one used in the ordinary formal theories of nuclear resonance reactions in which the decaying tails of various internal wave functions are neglected.

In the external region defined in this way, there are some regions where the internal

*) The reader is referred to a forthcoming paper of M. Kawai and the present author for detailed relations between these two formalisms.

wave functions of some pairs overlap with those of another pairs. In the example discussed above, the deuteron wave function overlaps with the wave function of the neutron or proton trapped within the range of the interaction with the target nucleus, while the neutron and the proton are outside the range of their mutual interaction. In other words, even in the external region there are sources of waves describing various pairs, and the strengths of these sources can be known only when the total wave function, i.e., the incident wave plus the scattered wave, is determined. From these sources there arise outgoing and incoming waves of various pairs. These outgoing waves give reaction amplitudes other than those arising from sources in the internal region, that is, those due to the formation and subsequent decay of the compound nucleus. On the other hand, these incoming

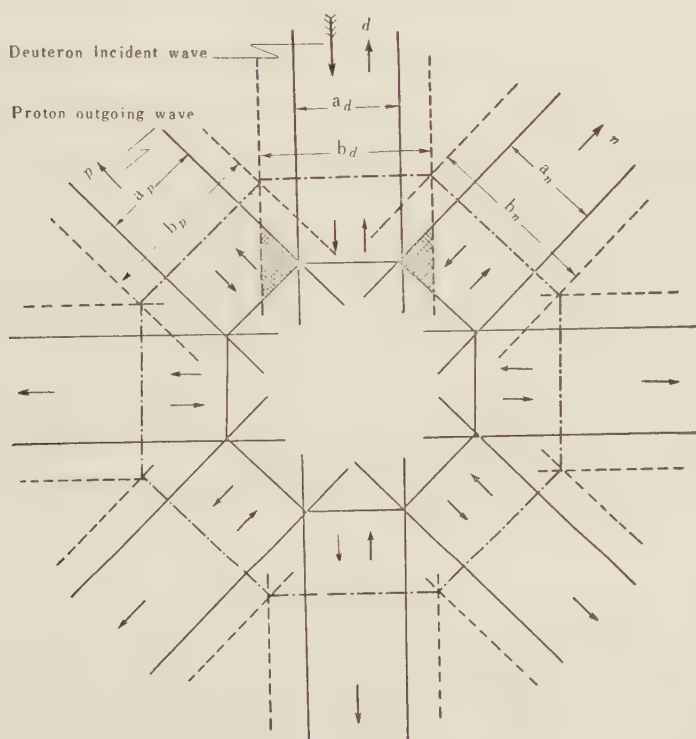


Fig. 1. Schematic representation of the configuration space. a_d , a_p , a_n are the ranges of the internal interactions acting to make up the pairs of the deuteron-the target nucleus, the proton-its residual nucleus, and the neutron-its residual nucleus, respectively. b_d , b_p , b_n are the extensions of the internal wave functions of these pairs. Along the "channels" denoted by d , p and n , we measure the relative distances between the two particles of these pairs. The polygon enclosed by the dot-and-dash line is the internal region described in the text, for which the ordinary formalisms of nuclear resonance reactions can be applied without correction. The polygon enclosed by the full line is the internal region to be used in this paper, where all the strong interactions of the system are acting effectively to form the compound state.

waves modify the boundary condition, which the scattered wave of the system as a function in the configuration space must satisfy on the boundary surface of the internal region, from the outgoing boundary condition in the case of no overlappings.

Our first approximation to be done is to give an approximate expression for the scattered wave in determining the strengths of the sources in the external region. We approximate also the boundary condition on the surface of the internal region. The incoming waves arising from the sources in the external region are taken as the "effective" incident waves, besides the incident plane wave describing the relative motion of the initial pair initiating the reaction process. We then use the collision matrix of Kapur and Peierls^{1,2} in order to relate the amplitudes of the outgoing waves to the amplitudes of the incident waves on the surface of the internal region, neglecting the overlappings of the internal wave functions on the surface.

Since this problem of the effect of the overlapping of the internal wave functions of pairs in the external region is not special in the deuteron stripping reaction, but common to many nuclear reactions such as (p, d) pick-up reaction, (n, p) direct reaction or nuclear reactions with which some loosely bounded particles concern, we shall formulate this problem in § 2 in a general way so as to be applicable also to these reactions. In § 3 we shall apply the result of § 2 to (d, p) reactions, showing that the outgoing proton waves arising from the sources in the external region, which are composed only of the incident plane wave, give us the Butler type angular distribution of emerging protons. We shall also discuss qualitatively the effects of the compound nucleus formation on the angular distribution of emerging protons. Some comments will be made on the relation between our present formulation and some of the theories hitherto presented. As in the formal theories of Kapur and Peierls^{1,2} and of Wigner^{1,3}, the antisymmetrization of the total wave function will not be performed, and all the interactions of the system will be assumed to be of nonexchange character.

§ 2. General formulation

In this section we consider the process of the nuclear reaction initiated by the collision of a pair of two particles. Particles, of course, mean composite ones as well as elementary ones. The initial pair of two particles will be denoted by s , while a pair of two particles in the reaction products will be denoted by t , including the initial one. For example, in the reaction



" s " denotes the pair of the initial two particles O^{16} and d , while " t " denotes any one of the pairs of the final two particles O^{16} and d , O^{17} and p , F^{17} and n , etc.

For a pair t , the total Hamiltonian of the system H can be split as

$$H = H_t + T_t + V_t. \quad (2 \cdot 1)$$

Here H_t is the internal Hamiltonian for the two particles of the pair t , T_t is the kinetic energy operator for the relative motion between them, and V_t is the potential energy of their mutual interaction. The product of the normalized internal wave functions of the two particles of the pair t will be denoted by $\chi_{t\tau}(i_t)$, so that $\chi_{t\tau}$ satisfies the equation

$$(H_t - \varepsilon_{t\tau}) \chi_{t\tau}(i_t) = 0. \quad (2.1a)$$

i_t represents all the internal coordinates of the two particles of the pair t , including those of their spins. τ determines the state of excitation of the pair t , $\varepsilon_{t\tau}$, together with the angular momentum resulting from the spins of the pair t .

We write the total wave function Ψ_s as

$$\Psi_s = \Phi_s + \psi_s^{sc}. \quad (2.2)$$

Φ_s is the incident plane wave of the pair s with unit flux, i.e.,

$$\Phi_s = v_s^{-1/2} e^{i\mathbf{k}_{s0} \cdot \mathbf{r}_s} \chi_{s0}(i_s), \quad (2.2a)$$

where v_s is the relative velocity between the two particles of the pair s , \mathbf{r}_s is the vector connecting the centers of mass of the two particles, and σ denotes the state of excitation of the pair s together with the angular momentum resulting from their spins. ψ_s^{sc} is the scattered wave describing the reaction process initiated by the collision of the pair s . If E is the energy of the system, it is not difficult to see from the Schrodinger equation for Ψ_s , $(E - H) \Psi_s = 0$, that ψ_s^{sc} satisfies the equation

$$(E - H_t - T_t) \psi_s^{sc} = V_s \Phi_s + V_t \psi_s^{sc}.$$

Multiplying by $\chi_{t\tau}^*$ and integrating with respect to the internal coordinates i_t , we have

$$(\nabla_t^2 + k_{t\tau}^2) R_{t\tau}(\rho_t) = 2M_t \hbar^{-2} \int d i_t \chi_{t\tau}^* (V_s \Phi_s + V_t \psi_s^{sc}), \quad (2.3)$$

with

$$R_{t\tau}(\rho_t) = \int d i_t \chi_{t\tau}^* \psi_s^{sc}. \quad (2.3a)$$

Here ρ_t is the vector connecting the centers of mass of the two particles of the pair t , ∇_t^2 is the Laplacian operator for the relative motion between the two particles of the pair t , and $\hbar^2 k_{t\tau}^2 / 2M_t = E - \varepsilon_{t\tau}$ is the available kinetic energy for the relative motion, $k_{t\tau}$ and M_t being the wave number and the reduced mass, respectively, associated with the relative motion.

$R_{t\tau}$ must satisfy the appropriate boundary conditions for the reaction. If $G_{t\tau}(\rho_t, \rho_t')$ is the Green function of the operator $(\nabla_t^2 + k_{t\tau}^2)$ which satisfies the outgoing boundary condition at $|\rho_t| \rightarrow \infty$, namely $(-4\pi|\rho_t - \rho_t'|)^{-1} e^{i k_{t\tau} |\rho_t - \rho_t'|}$, $R_{t\tau}$ can be expressed as

$$R_{t\tau}(\rho_t) = 2M_t \hbar^{-2} \int d \rho_t' d i_t G_{t\tau}(\rho_t, \rho_t') \chi_{t\tau}^* (V_s \Phi_s + V_t \psi_s^{sc}). \quad (2.4)$$

If $R_{t\tau}(\rho_t)$ is expanded in terms of the normalized spherical harmonics $Y_{l_t m_t}(\Omega_t)$ with its coefficient $R_{l_t m_t}(\rho_t)$, where ρ_t and Ω_t are the length and the direction of the vector ρ_t respectively,

$$R_{t\tau}(\rho_t) = \sum_{l_t m_t} R_{t\tau l_t m_t}(\rho_t) Y_{l_t m_t}(\Omega_t), \quad (2.5)$$

we have from (2.4) with the usual expansion of $G_{t\tau}$ in the spherical harmonics,

$$R_{t\tau l_t m_t} = (2M_t k_{t\tau} / i\hbar^3) \{ d\rho'_t \, di_t \, j_{l_t}(k_{t\tau} \rho_{t<}) h_{l_t}^{(1)}(k_{t\tau} \rho_{t>}) \\ \cdot Y_{l_t m_t}^* \chi_{t\tau}^* (V_s \Phi_s + V_t \phi_s^{sc}). \quad (2.6)$$

In this equation, j_l and $h_l^{(1)}$ are the spherical Bessel and Hankel functions of the first kind of order l ,* and $\rho_{t<}$ denotes the smaller of ρ_t and ρ'_t , while $\rho_{t>}$ the larger of them. Let us define the functions $\varphi_{t\tau l_t m_t}^{(+)}$, $\varphi_{t\tau l_t m_t}^{(-)}$, and $\varphi_{t\tau l_t m_t}$ as

$$\begin{cases} \varphi_{t\tau l_t m_t}^{(+),(-)} = (\hbar/M_t k_{t\tau})^{-1/2} h_{l_t}^{(1),(2)}(k_{t\tau} \rho_t) Y_{l_t m_t}(\Omega_t) \chi_{t\tau}(i_t), \\ \varphi_{t\tau l_t m_t} = (\hbar/M_t k_{t\tau})^{-1/2} j_{l_t}(k_{t\tau} \rho_t) Y_{l_t m_t}(\Omega_t) \chi_{t\tau}(i_t). \end{cases} \quad (2.7)$$

$\varphi_{t\tau l_t m_t}^{(+)}$ has unit outgoing flux over the full solid angle, while $\varphi_{t\tau l_t m_t}^{(-)}$ has unit incoming flux. $R_{t\tau l_t m_t}$ becomes

$$R_{t\tau l_t m_t} \chi_{t\tau} Y_{l_t m_t} = (2/i\hbar) \varphi_{t\tau l_t m_t} \int_{(\rho_t > \rho_t)} d\rho'_t \, di_t \, \varphi_{t\tau l_t m_t}^{(-)*} (V_s \Phi_s + V_t \phi_s^{sc}) \\ + (2/i\hbar) \varphi_{t\tau l_t m_t}^{(+)} \int_{(\rho_t' < \rho_t)} d\rho'_t \, di_t \, \varphi_{t\tau l_t m_t}^* (V_s \Phi_s + V_t \phi_s^{sc}). \quad (2.8)$$

At the channel radius $\rho_t = a_t$, $R_{t\tau l_t m_t}(\rho_t)$ has both the outgoing and standing waves. (At $\rho_t = a_t$ the amplitude of the standing wave $\varphi_{t\tau l_t m_t}$ does not vanish, because the channel radius a_t is defined in the same way as in the ordinary formal theories of nuclear resonance reactions in which the decaying tails of the internal wave functions of the two particles of the pair t are entirely neglected. If the overlappings of the internal wave functions in the external region are taken into account, there remain some sources of the pair t even for $\rho_t > a_t$). We denote the amplitudes of these outgoing and standing waves at $\rho_t = a_t$ by $u_{t\tau l_t m_t}$ and $I_{t\tau l_t m_t}$ respectively. $u_{t\tau l_t m_t}$ is given by

$$u_{t\tau l_t m_t} = (2/i\hbar) \int_{(\rho_t = a_t)} d\rho'_t \, di_t \, \varphi_{t\tau l_t m_t}^* (V_s \Phi_s + V_t \phi_s^{sc}), \quad (2.9a)$$

while for $I_{t\tau l_t m_t}$ we have

$$I_{t\tau l_t m_t} = (2/i\hbar) \int_{(\rho_t > a_t)} d\rho'_t \, di_t \, \varphi_{t\tau l_t m_t}^{(-)*} (V_s \Phi_s + V_t \phi_s^{sc}), \quad (2.9b)$$

where we have replaced ρ'_t , the variable of integration, by ρ_t without confusion. For the pair s , we have, besides $I_{s0 l_s m_s}$, the amplitude $I_{s0 l_s 0}^{(i)}$ of the l_s -th partial wave $\varphi_{s0 l_s 0}$ in the incident plane wave Φ_s :

$$I_{s0 l_s 0}^{(i)} = k_{s0}^{-1} i^{l_s} \sqrt{4\pi(2l_s+1)}. \quad (2.10)$$

* As is obvious, when the two particles of the pair t have electric charges, j_l and $h_l^{(1)}$ must be replaced by their corresponding Coulomb wave functions. We neglect polarizations of particles in Coulomb fields. This we do because we are interested here in the connection between the ordinary resonance theory and the effect of the overlapping in the external region. In some actual cases this effect of polarization seems to be important.

In the overlapping part of the external region, there are also both outgoing and standing waves, whose amplitudes change as we go from the surface of the internal region to the non-overlapping region. Our first step of approximation is, as is indicated in the introduction, to give an approximate expression for ψ_s^{sc} in the defining integral of $I_{\tau l m_l}$. As a first approximation, we neglect the amplitudes of standing waves and neglect also the change of the amplitudes of outgoing waves. Thus in order to calculate $I_{\tau l m_l}$, we put ψ_s^{sc} in the external region as

$$\psi_s^{sc} = \sum_{l' \tau' l'_l m'_l} U_{l' \tau' l'_l m'_l} \varphi_{l' \tau' l'_l m'_l}^{(+)}, \quad (2.11)$$

where $U_{\tau l m_l}$ is the amplitude of $\varphi_{\tau l m_l}^{(+)}$ in the non-overlapping region, that is to say in the "wave-zone", which we are trying to obtain.

Inserting this expression into (2.9b) we have

$$I_{\tau l m_l} = I_{\tau l m_l}^{(d)} + \sum_{l' \tau' l'_l m'_l} \alpha_{\tau l m_l} \varphi_{l' \tau' l'_l m'_l}^{(+)} U_{l' \tau' l'_l m'_l} \quad (2.12a)$$

where

$$I_{\tau l m_l}^{(d)} = (2/i\hbar) \int d\tau \varphi_{\tau l m_l}^{(-)*} V_s \Phi_s, \quad (\rho_l > a_l) \quad (2.12b)$$

$$\alpha_{\tau l m_l; l' \tau' l'_l m'_l} = (2/i\hbar) \int d\tau \varphi_{\tau l m_l}^{(-)*} V_l \varphi_{l' \tau' l'_l m'_l}^{(+)} \quad (\rho_l > a_l) \quad (2.12c)$$

and $\int d\tau$ denotes the integration with respect to all the coordinates of the system in the center of mass coordinate system (including the summation with respect to all the spin-coordinates).

The relation connecting $U_{\tau l m_l}$ and $u_{\tau l m_l}$ can be obtained from (2.8). Noting that $U_{\tau l m_l}$ is the amplitude of $\varphi_{\tau l m_l}^{(+)}$ at the non-overlapping region, we let ρ_l in (2.8) be so large that the source of the pair t , i.e. $\{d_l, \chi_t^*(V, \Phi, +V_l \psi_s^{sc})\}$ vanishes. Then the first term of the right hand side of (2.8) becomes zero, leaving us only the outgoing wave $\varphi_{\tau l m_l}^{(+)}$, whose amplitude is given by

$$U_{\tau l m_l} = (2/i\hbar) \int d\rho_l' d_l \varphi_{\tau l m_l}^* (V_s \Phi_s + V_l \psi_s^{sc}).$$

Replacing ρ_l' by ρ_l without confusion, and using (2.9a), this can be rewritten as

$$U_{\tau l m_l} = u_{\tau l m_l} + (2/i\hbar) \int d\tau \varphi_{\tau l m_l}^* (V_s \Phi_s + V_l \psi_s^{sc}) \quad (\rho_l > a_l) \quad (2.13)$$

$u_{\tau l m_l}$ has been defined to be the amplitude of $\varphi_{\tau l m_l}^{(+)}$ arising from sources in the region $\rho_l < a_l$. This region includes not only all the parts of the internal region but also some parts of the external region. We therefore divide $u_{\tau l m_l}$ into two parts. One is the amplitude of $\varphi_{\tau l m_l}^{(+)}$ due to the sources in the internal region, and the other is that arising from the sources in the external region with $\rho_l < a_l$. The latter one gives us, together with the second term of (2.13), the part of the amplitude of $\varphi_{\tau l m_l}^{(+)}$ due to the sources only in the external region. Thus we have, again approximating ψ_s^{sc} by (2.11)

$$U_{\tau l m_l} = u_{\tau l m_l}^{(c)} + u_{\tau l m_l}^{(d)} + \sum_{l' \tau' l'_l m'_l} \beta_{\tau l m_l; l' \tau' l'_l m'_l} U_{l' \tau' l'_l m'_l}, \quad (2.14a)$$

with

$$u_{l\tau l_i m_i}^{(c)} = (2/i\hbar) \int_{(int)} d\tau \varphi_{l\tau l_i m_i} * (V_s \Phi_s + V_l \phi_s^{sc}), \quad (2 \cdot 14b)$$

$$u_{l\tau l_i m_i}^{(d)} = (2/i\hbar) \int_{(ext)} d\tau \varphi_{l\tau l_i m_i} * V_s \Phi_s, \quad (2 \cdot 14c)$$

and

$$\beta_{l\tau l_i m_i; l'\tau' l'_i m'_i} = (2/i\hbar) \int_{(ext)} d\tau \varphi_{l\tau l_i m_i} * V_l \varphi_{l'\tau' l'_i m'_i}^{(+)}, \quad (2 \cdot 14d)$$

where $\int_{(int)} d\tau$ denotes the integration with respect to all the coordinates as in (2.12), but is extended only over the internal region, while $\int_{(ext)} d\tau$ is similar one extended only over the external region. $u_{l\tau l_i m_i}^{(c)}$ will be called the amplitude due to the compound formation, and $u_{l\tau l_i m_i}^{(d)}$ the direct amplitude due to the overlapping in the external region.

Our next step is to treat approximately the boundary condition imposed on the total wave function of the system on the surface of the internal region. If on the surface of the internal region $\chi_{l\tau} Y_{l_i m_i}$'s form an orthonormal complete set of functions with l , τ , l_i and m_i as the index of member, it is an exact procedure of matching the boundary condition to join smoothly each $R_{l\tau l_i m_i}$ at $r_i = a_i$. We would have this orthonormality if we were taking large enough an internal region. In such a case, however, as is mentioned in the introduction, the internal region would loose its close relationship to the notion of the compound nucleus. In the theory of Kapur and Peierls¹³, this lack of the orthonormality due to the overlappings of the internal wave functions on the surface of the internal region was entirely neglected, as well as were the effects of the overlappings in the external region.

In our present approximation, we are also neglecting this lack of the orthonormality of $\chi_{l\tau} Y_{l_i m_i}$'s on the surface of the internal region. However, the neglect of the lack of the orthonormality in our case is expected to be less inadequate than in the case of Kapur and Peierls, because we are taking into account the effects of the "multiple processes" due to the overlappings in the external region. Thus the matching of the boundary condition, which must be satisfied by the total wave function of the system on the surface of the internal region, will be treated approximately by taking $I_{l\tau l_i m_i}^{(sc)}$, $I_{l\tau l_i m_i}$ and $u_{l\tau l_i m_i}^{(c)}$ as the amplitudes of the standing and outgoing waves in the theory of Kapur and Peierls. In their theory, the collision matrix $S_{l\tau l_i m_i; l'\tau' l'_i m'_i}$, which gives the amplitude of the outgoing wave $\varphi_{l\tau l_i m_i}^{(+)}$ when only the standing wave of unit amplitude $\varphi_{l'\tau' l'_i m'_i}$ is incident upon the internal region, has been expressed as

$$S_{l\tau l_i m_i; l'\tau' l'_i m'_i} = (i/2) \{1 - h_{l_i}^{(1)*}(k_{l\tau} a_i) / h_{l_i}^{(1)}(k_{l\tau} a_i)\} \delta_{ll'} \delta_{\tau\tau'} \delta_{l_i l'_i} \delta_{m_i m'_i} \\ + \frac{1}{2} \sum_n \frac{u_{l\tau l_i m_i}^{(n)} u_{l'\tau' l'_i m'_i}^{(n)}}{N_n (W_n - E)}.$$

Here, W_n is the (complex) resonance energy of the n -th level of the compound nucleus, and $u_{l\tau l_i m_i}^{(n)}$ is proportional to the amplitude of $\varphi_{l\tau l_i m_i}^{(+)}$ which the n -th level has on the surface of the internal region. The first term of this expression is the well-known amplitude for the so-called potential scattering.

In our present case, there are various incident standing waves of amplitudes given by $I^{(s)}$ and I . The amplitude of the outgoing wave $\varphi_{t\tau l_t m_t}^{(+)}$ arising from the sources in the internal region, $u_{t\tau l_t m_t}^{(c)}$, is therefore given by

$$u_t^{(c)} = \sum_{l_s} S_{t,s} I_s^{(s)} + \sum_{l'} S_{t,l'} I_{l'}^{(t)} + \sum_{l'} \sum_{l''} S_{t,l'} \alpha_{l',l''} U_{l''}, \quad (2.15)$$

abbreviating the set of suffices (t, τ, l_t, m_t) , etc., simply as t , etc..

Considering $U, u^{(c)}, u^{(t)}, I^{(s)}$ and $I^{(t)}$ as vectors with components $U_t, u_t^{(c)}, u_t^{(t)}, I_s^{(s)}$ and $I_t^{(t)}$ respectively, and S, α and β as matrices with their respective elements $S_{t,l'}, \alpha_{t,l'}$ and $\beta_{t,l'}$, we have from (2.14a)

$$U = (1 - \beta)^{-1} (u^{(t)} + u^{(c)}). \quad (2.16a)$$

Inserting this into (2.15), we have for $u^{(c)}$

$$u^{(c)} = \{1 - S\alpha(1 - \beta)^{-1}\}^{-1} S \{I^{(s)} + I^{(t)} + \alpha(1 - \beta)^{-1} u^{(t)}\}. \quad (2.16b)$$

The differential cross section in the center of mass coordinate system for the reaction, in which the two particles of the pair s , their relative motion being described by $v_s^{-1/2} \exp[ik_{s\sigma} o_s]$, react with each other and the two particles of the pair t emerge, is given by

$$d\sigma_{t\tau; s\sigma} = |\sum_{l_t m_t} i^{-l_t} U_{t\tau l_t m_t} Y_{l_t m_t}(\varrho_t)|^2 d\varrho_t. \quad (2.18)$$

(In order to compare with experiments, this must, of course, be averaged over the orientations of the initial spins, and be summed over those of the final spins with a usual procedure).

§ 3. The deuteron stripping

In this section, we apply the result of the preceding section to (d, p) reactions. We are interested in the differential cross section of the emerging protons in the reaction initiated by the collision between a deuteron and a target nucleus. The pair of the incident deuteron and the target nucleus will be denoted by d , instead of s in the preceding section, while that of the proton and its residual nucleus will be denoted by p . The states of excitation of the pairs d and p will be denoted by their corresponding Greek letters δ and π , respectively, together with the angular momenta resulting from their spins. Thus, $\chi_{d\delta}$ describes the internal motions of the deuteron and the target nucleus in its ground state, while $\chi_{p\pi}$ represents the internal motion of the residual nucleus in its ground or some excited state together with the orientation of the intrinsic spin of the proton.

For the amplitude of the outgoing proton wave with its orbital angular momentum l_p, m_p relative to the residual nucleus in some excitation, we have from (2.16)

$$U_{p\pi l_p m_p} = [(1 - \beta)^{-1} \{u^{(t)} + (1 - S\alpha(1 - \beta)^{-1})^{-1} S (I^{(s)} + I^{(t)} + \alpha(1 - \beta)^{-1} u^{(t)})\}]_{p\pi l_p m_p}. \quad (3.1)$$

The direct amplitude $u^{(t)}$ is from (2.14c)

$$u_{p\pi l_p m_p}^{(t)} = (2/i\hbar) \int d\tau \varphi_{p\pi l_p m_p}^* V_d \Phi_d. \quad (3.2)$$

ϕ_d is the plane wave of the deuteron incident upon the target nucleus,

$$\phi_d = v_d^{-1/2} \exp[i\mathbf{k}_d \cdot \mathbf{r}_d] \chi_{d\delta}(i_d), \quad (3.3)$$

where we have neglected, for simplicity, the Coulomb interaction between the deuteron and the target nucleus. V_d is the interaction between the deuteron and the target nucleus, and is given by

$$V_d = V_{p\pi} + V_{n\pi}$$

where $V_{p\pi}$ is the interaction between the proton and the target nucleus, while $V_{n\pi}$ is that between the neutron and the target nucleus.

The wave function which describes the emerging proton can be obtained by summing up the outgoing wave $\varphi_{p\pi l_p m_p}^{(+)}$ over l_p, m_p with amplitude $U_{p\pi l_p m_p}$. If we use for $U_{p\pi l_p m_p}$ only the direct amplitude $u_{p\pi l_p m_p}^{(d)}$ it is easily seen that this is nothing but the outgoing wave function arising from the sources in the external region with strength $2M_p \hbar^{-2} \{d_{ip} \chi_{p\pi}^* V_d \phi_d\}$. Thus if we write this wave function evaluated at $\rho_p \rightarrow \infty$ as $v_p^{-1/2} \rho_p^{-1} e^{ik_{p\pi} \rho_p} \int_{p\pi}^{(d)} \chi_{p\pi}$, we have for $f_{p\pi}^{(d)}(\Omega_p)$, neglecting also the Coulomb interaction between the proton and the residual nucleus,

$$f_{p\pi}^{(d)}(\Omega_p) = (-M_p/2\pi\hbar^2) (v_{p\pi}/v_d)^{1/2} \int_{(e,t)} d\tau e^{-i\mathbf{k}_{p\pi} \cdot \mathbf{r}_p} \chi_{p\pi}^* V_d e^{i\mathbf{k}_d \cdot \mathbf{r}_d} \chi_{d\delta}, \quad (3.4)$$

where $\mathbf{k}_{p\pi}$ is the vector with length $k_{p\pi}$ and direction Ω_p . The angular distribution of emerging protons which arises from the direct amplitude alone is therefore

$$d\sigma_{p\pi, d\delta} = \frac{M_p M_d k_{p\pi}}{(2\pi\hbar^2)^2 k_d} \left| \int_{(e,t)} d\tau e^{-i\mathbf{k}_{p\pi} \cdot \mathbf{r}_p} \chi_{p\pi}^* (V_{p\pi} + V_{n\pi}) e^{i\mathbf{k}_d \cdot \mathbf{r}_d} \chi_{d\delta} \right|^2, \quad (3.5)$$

where M_p and M_d are the reduced masses of the proton and the deuteron, respectively. This result has very analogous appearance to that of the Born approximation, excepting the range of the integration with respect to all the coordinates of the system. If the integration were carried over the full intervals, it would be nothing but the Born approximation. In our result, however, it must be carried only over the external region. Since $V_{p\pi}$ contributes through the decaying tail of the internal wave function of the residual nucleus into the neutron channel beyond a_n (where a_n is the channel radius for the neutron), we can safely neglect the contribution from $V_{p\pi}$. Then our result becomes much similar to the Born approximation used by Bhatia *et al.*⁽²⁾ and by Ditch and French⁽³⁾.

If we denote the vectors connecting the center of mass of the target nucleus and the proton with \mathbf{r}_p , and the neutron with \mathbf{r}_n , the integral in (3.5) becomes

$$\begin{aligned} I &\equiv \int_{(e,t)} d\tau e^{-i\mathbf{k}_{p\pi} \cdot \mathbf{r}_p} \chi_{p\pi}^* V_{p\pi} e^{i\mathbf{k}_d \cdot \mathbf{r}_d} \chi_{d\delta} \\ &= \int d\mathbf{r}_p \int d\mathbf{r}_n \left\{ d\tilde{\zeta} e^{-i\mathbf{k}_{p\pi} \cdot \{\mathbf{r}_p - (M/M_f)\mathbf{r}_n\}} \chi_{p\pi}^*(\tilde{\zeta}, \mathbf{r}_n) V_{n\pi} \chi_0(\tilde{\zeta}) w_d(\mathbf{r}_p - \mathbf{r}_n) e^{i\mathbf{k}_d \cdot (\mathbf{r}_p + \mathbf{r}_n)} / 2 \right. \\ &\quad \left. (\mathbf{r}_p > a) (\mathbf{r}_n \leq a_0) \right\} \end{aligned} \quad (3.6)$$

M_f and M are the masses of the residual nucleus and a nucleon respectively. w_d is the internal wave function of the deuteron, while χ_0 represents the internal wave function of

the target nucleus, together with the spin functions of the system, ξ being the internal coordinates of the target nucleus and all the spin coordinates of the system. We have neglected the D -state component of the deuteron wave function. a_0 is the range of the interaction $V_{n\bar{c}}$ between the neutron and the target nucleus, while a represents the range of the interaction between the proton and the residual nucleus measured from the center of mass of the target nucleus. Denoting by $H_{\bar{c}}$ the Hamiltonian for the internal motion of the target nucleus together with that for all the spin freedoms, we have for $r_p > a$ and $r_n \leq a_0$,

$$\left\{ \begin{aligned} & \left[H_{\bar{c}} + \frac{-\hbar^2}{2M} \nabla_p^2 + \frac{-\hbar^2}{2M} \nabla_n^2 + \frac{-\hbar^2}{2M_0} \nabla_d^2 - E \right] \chi_0(\xi) \psi_d(\mathbf{r}_p, \mathbf{r}_n) = 0, \\ & \left[H_{\bar{c}} + V_{n\bar{c}} + \frac{-\hbar^2}{2M} \nabla_p^2 + \frac{-\hbar^2}{2M} \nabla_n^2 + \frac{-\hbar^2}{2M_0} \nabla_d^2 - E \right] e^{-ik_{p\pi}\{\mathbf{r}_p - (M/M_f)\mathbf{r}_n\}} \chi_{p\pi}^*(\xi, \mathbf{r}_n) = 0, \end{aligned} \right.$$

where $\psi_d = w_d(\mathbf{r}_p - \mathbf{r}_n) e^{ik_d(\mathbf{r}_p + \mathbf{r}_n)/2}$ and ∇_p^2 , ∇_n^2 and ∇_d^2 are the Laplacians for \mathbf{r}_p , \mathbf{r}_n and $\rho_d = (\mathbf{r}_p + \mathbf{r}_n)/2$, respectively, and M_0 is the mass of the target nucleus. E is the energy of the system in the center of mass coordinate. From these equations it will be easily shown that

$$\begin{aligned} & \int d\xi e^{-ik_{p\pi}\{\mathbf{r}_p - (M/M_f)\mathbf{r}_n\}} \chi_{p\pi}^* V_{n\bar{c}} \chi_0 \psi_d \\ & = (\hbar^2/2M) \{ \psi_d (\nabla_p^2 + \nabla_n^2) e^{-ik_{p\pi}\mathbf{r}_p} \zeta^* - \zeta^* e^{-ik_{p\pi}\mathbf{r}_p} (\nabla_p^2 + \nabla_n^2) \psi_d \}, \end{aligned} \quad (3.7a)$$

with

$$\zeta(\mathbf{r}_n) = \int d\xi e^{-i(M/M_f)k_{p\pi}\mathbf{r}_n} \chi_{p\pi}(\xi, \mathbf{r}_n) \chi_0^*(\xi). \quad (3.7b)$$

Here we have neglected contributions from the terms containing $\hbar^2 \nabla_d^2 / M_0$. $\zeta(\mathbf{r}_n)$ describes the motion of the neutron in the residual nucleus viewed from the target nucleus. Inserting this result into (3.6), it becomes

$$\begin{aligned} I = & -\frac{\hbar^2}{2M} \int_{(\mathbf{r}_n \leq a_0)} d\mathbf{r}_n \zeta^* \cdot \int a^2 d\Omega_p \left[\psi_d \frac{\partial}{\partial \mathbf{r}_p} e^{-ik_{p\pi}\mathbf{r}_p} - e^{-ik_{p\pi}\mathbf{r}_p} \frac{\partial}{\partial \mathbf{r}_p} \psi_d \right]_a \\ & + \frac{\hbar^2}{2M} \int d\mathbf{k} G(\mathbf{k}) g(\mathbf{k}_{p\pi} - (1/2)\mathbf{k}_d + \mathbf{k}) \int a_0^2 d\Omega_n \left[e^{i(\mathbf{k}_d/2 + \mathbf{k})\mathbf{r}_n} \frac{\partial}{\partial \mathbf{r}_n} \zeta^* - \zeta^* \frac{\partial}{\partial \mathbf{r}_n} e^{i(\mathbf{k}_d/2 + \mathbf{k})\mathbf{r}_n} \right]_{a_0} \end{aligned} \quad (3.8a)$$

where we have put G and g as

$$\left\{ \begin{aligned} & w_d(\mathbf{r}_p - \mathbf{r}_n) = \int d\mathbf{k} G(\mathbf{k}) e^{i\mathbf{k}(\mathbf{r}_p - \mathbf{r}_n)} \\ & g(\mathbf{q}) = \int_{(\mathbf{r}_p > a)} d\mathbf{r}_p e^{-i\mathbf{q}\mathbf{r}_p}. \end{aligned} \right. \quad (3.8b)$$

The first term of (3.8a) includes an integration with respect to \mathbf{r}_n over the interval $r_n \leq a_0$, where $\zeta(\mathbf{r}_n)$ oscillates rapidly on account of the strong interaction between the neutron and the target nucleus. We can, therefore, reasonably neglect this term, as is done in Daitch and French's treatment with the Born approximation and also implicitly

in Butler's treatment. The second term gives to (3.5) a factor which is very analogous to Butler's factor determining the angular distribution of stripped-off protons. Indeed, if the integration with respect to \mathbf{r}_p were carried over the full interval, it would give the same result with Butler's. In our present case, however, this integration must be carried only over the interval $r_p > a$. $g(\mathbf{q})$ can be rewritten as

$$g(\mathbf{q}) = (2\pi)^3 \delta(\mathbf{q}) - 4\pi a^3 j_1(qa)/qa, \quad (3.8c)$$

where $j_1(qa)$ is the spherical Bessel function of order one. We thus have for I

$$I = \frac{\hbar^2 a_0^2}{2M} (2\pi)^3 G(\mathbf{k}_{p\pi} - \mathbf{k}_d/2) \int d\Omega_n \left[e^{i(\mathbf{k}_d - \mathbf{k}_{p\pi}) \cdot \mathbf{r}_n} \frac{\partial}{\partial \mathbf{r}_n} \cdot \mathbf{r}_n^* - \mathbf{r}_n^* \frac{\partial}{\partial \mathbf{r}_n} e^{i(\mathbf{k}_d - \mathbf{k}_{p\pi}) \cdot \mathbf{r}_n} \right]_{r_n=a} \\ - \frac{\hbar^2 a_0^2}{2M} (4\pi a^3) \int d\Omega_n \left[e^{i(\mathbf{k}_d - \mathbf{k}_{p\pi}) \cdot \mathbf{r}_n} \gamma_i(\mathbf{r}_n) \frac{\partial}{\partial \mathbf{r}_n} \cdot \mathbf{r}_n^* - \mathbf{r}_n^* \frac{\partial}{\partial \mathbf{r}_n} e^{i(\mathbf{k}_d - \mathbf{k}_{p\pi}) \cdot \mathbf{r}_n} \gamma_i(\mathbf{r}_n) \right]_{a_0} \quad (3.9a)$$

with

$$\gamma_i(\mathbf{r}_n) = \sum_{m=0}^{\infty} \frac{(-1)^m 2(m+1)}{(2m+3)!} (a^2 \nabla_n^2)^m \{ e^{i(\mathbf{k}_{p\pi} - \mathbf{k}_d/2) \cdot \mathbf{r}_n} w_d(\mathbf{r}_n) \}. \quad (3.9b)$$

The first term of this result gives just the Butler factor. The second term gives a correction to Butler's factor arisen from our present formulation. In his treatment, Butler has ignored the possibility of both the neutron and the proton being simultaneously within the range of their interactions with the target nucleus. Though he has considered that the deuteron stripping (d, p) reaction occurs when only the neutron interacts with the target nucleus, in his actual calculation, neglecting entirely the interaction between the proton and the target nucleus, the total wave function of the system has been analysed in terms of the Fourier components of the proton plane wave, and each Fourier component has been adjusted to join the interior and the exterior wave functions of the neutron at the nuclear surface. The contributions from the configurations in which both the proton and the neutron are within the range of their interactions with the target nucleus are therefore approximated by his stripping amplitude. In our present formulation, these contributions are included in the amplitude $u^{(1)}$ due to the compound nucleus formation. Though this correction term tends to decrease the peaked forward angular distribution, this correction does not dominate if the radius a is not large compared with the wave length $|\mathbf{k}_{p\pi} - \mathbf{k}_d/2|^{-1}$ and the extension of the deuteron internal wave function. For heavy nuclei, however, this correction may play some important roles in determining the angular distribution of the stripped-off protons.

It is a well-known fact that the Born approximation $\{ d\tau e^{-i\mathbf{k}_{i\tau} \cdot \mathbf{r}_i} \chi_{i\tau}^* V_i e^{i\mathbf{k}_{s0} \cdot \mathbf{r}_s} \chi_{s0}$ and $\{ d\tau e^{-i\mathbf{k}_{i\tau} \cdot \mathbf{r}_i} \chi_{i\tau}^* V_i e^{i\mathbf{k}_{s0} \cdot \mathbf{r}_s} \chi_{s0}$ give an identical result for the differential cross section $d\sigma_{i\tau, s0}/d\Omega_i$. In some of the theories on the deuteron stripping reactions, therefore, the angular distributions of stripped-off nucleons have been treated in terms of the final interaction instead of the initial interaction. In our present case, the integration $\{ d\tau$ is not extended over the full intervals of all the coordinates of the system but only over the external region. As is easily seen from the Schrodinger equations satisfied by $e^{-i\mathbf{k}_{i\tau} \cdot \mathbf{r}_i} \chi_{i\tau}^*$ and $e^{-i\mathbf{k}_{s0} \cdot \mathbf{r}_s} \chi_{s0}$, we have with the aid of Green's theorem,

$$\begin{aligned} & \int_{(c,t')} d\tau \left\{ e^{-ik_t \tau p_t} \chi_{t\tau}^* V_s e^{ik_s \sigma p_s} \chi_{s\sigma} - e^{-ik_t \tau p_t} \chi_{t\tau}^* V_t e^{ik_s \sigma p_s} \chi_{s\sigma} \right\} \\ &= - \sum_{t'} \frac{\hbar^2}{2M_{t'}} \int dS_{t'} \left\{ e^{-ik_t \tau p_t} \chi_{t\tau}^* \frac{\partial}{\partial \rho_{t'}} e^{ik_s \sigma p_s} \chi_{s\sigma} - e^{ik_s \sigma p_s} \chi_{s\sigma} \frac{\partial}{\partial \rho_{t'}} e^{-ik_t \tau p_t} \chi_{t\tau}^* \right\}, \quad (3.10) \end{aligned}$$

where $dS_{t'}$ is the element of the part of the surface of the internal region corresponding to the pair t' . We neglect this surface integral from a similar reasoning to the one used in the discussion about the orthonormality of $\chi_{t\tau} Y_{l_t m_t}$'s on the surface in the preceding section, with the result

$$\int_{(c,t')} d\tau e^{-ik_t \tau p_t} \chi_{t\tau}^* V_t e^{ik_s \sigma p_s} \chi_{s\sigma} = \int_{(c,t')} d\tau e^{-ik_t \tau p_t} \chi_{t\tau}^* V_s e^{ik_s \sigma p_s} \chi_{s\sigma}.$$

This gives us, if we expand $e^{-ik_t \tau p_t}$ in partial waves, for $u^{(t)}$ instead of (2.14c)

$$u_{t\tau l_t m_t}^{(t)} = (2/i\hbar) \int_{(c,t')} d\tau \varphi_{t\tau l_t m_t}^* V_t \Phi_s, \quad (2.14c')$$

in the same degree of accuracy as (2.14c). Applying this result to the (d, p) reaction, it gives a very similar expression to the Born approximations used by Gerjuoy⁹⁾ and Tobocman¹¹⁾.

Using this result, the integral I defined in (3.6) is rewritten as

$$I = \int_{(r_p > a)} d\mathbf{r}_p \int_{(r_n > a_0)} d\mathbf{r}_n \int d\zeta e^{-ik_{p\pi} \{r_p - (M/M_f) r_n\}} \chi_{p\pi}^* V_{np} \chi_0 \psi_d, \quad (3.6')$$

where use has been made of $V_p = V_{p\pi} + V_{np}$, V_{np} being the interaction between the proton and the neutron, and we have neglected as in (3.6) contributions from $V_{p\pi}$. The integration with respect to \mathbf{r}_n must be carried over the interval $r_n > a_0$, since those configurations in which the neutron is within the range of its interaction with both the proton and the target nucleus belong to the internal region. Using the Schrodinger equations which $e^{-ik_{p\pi} \{r_p - (M/M_f) r_n\}} \chi_{p\pi}^*$ and $\chi_0 \psi_d$ satisfy for $r_p > a$ and $r_n > a_0$, we have with a similar procedure as used in the derivation of (3.7),

$$I = - \frac{\hbar^2}{2M} \int d\mathbf{r}_p \int_{(r_p > a)} d\mathbf{r}_n \left\{ \psi_d (\nabla_n^2 + V_p^2) e^{-ik_{p\pi} r_p} \zeta^* - e^{-ik_{p\pi} r_p} \zeta^* (\nabla_n^2 + V_p^2) \psi_d \right\}. \quad (3.7')$$

Introducing the Fourier components G and g , this becomes

$$\begin{aligned} I &= \frac{\hbar^2}{2M} \int_{(r_p > a_0)} d\mathbf{r}_n \zeta^* \int a^2 d\Omega_p \left[\psi_d \frac{\partial}{\partial r_p} e^{-ik_{p\pi} r_p} - e^{-ik_{p\pi} r_p} \frac{\partial}{\partial r_p} \psi_d \right] \\ &+ \frac{\hbar^2}{2M} \int d\mathbf{k} G(\mathbf{k}) g(\mathbf{k}_{p\pi} - (1/2)\mathbf{k}_d + \mathbf{k}) \int_{a_0}^{\infty} d\Omega_n \left[e^{i(\mathbf{k}_d/2 + \mathbf{k}) r_n} \frac{\partial}{\partial r_n} \zeta^* - \zeta^* \frac{\partial}{\partial r_n} e^{i(\mathbf{k}_d/2 + \mathbf{k}) r_n} \right]_{a_0}. \end{aligned} \quad (3.8a')$$

The first term of this equation contributes through an integration of the decaying tail of ζ^* beyond a_0 , and is therefore of the same order of magnitude as the surface integral appeared in (3.10). We can neglect this term consistently with our previous approximations. Then (3.8a') gives a result for I identical with (3.8a) which has been obtained

from the treatment in terms of the initial interaction.

In Gerjuoy's derivation of the Butler factor, the integration with respect to \mathbf{r}_p in (3.6') was carried over its full interval, though the target nucleus was taken as a potential field acting on the neutron and the proton. Tobocman used (3.6') without the restriction $r_n > a_0$, as a simplest example of his treatment with a modified wave method. In order to cut down the integration with respect to \mathbf{r}_n at a definite radius, he has assumed the range of the interaction between the proton and the neutron to be zero, thereby obtaining the Butler factor. As is seen by putting the range of V_{np} to be zero in Gerjuoy's expression for the reaction amplitude, Tobocman's one with the assumption of the zero range interaction becomes identical with Gerjuoy's. The final result of Gerjuoy's treatment is, however, independent of the range of the interaction between the proton and the neutron. This is the reason why Tobocman's treatment with the restriction $r_p > a$ resulted in the Butler factor, without the correction due to the restriction for the proton to be outside the range of the interaction with the target nucleus which we have discussed above.

In this connection it may be noted that in our present formulation the restriction $r_n > a_0$ arises naturally without the assumption of the zero range interaction between the proton and the neutron. The channel radius a_p has been defined as the smallest one which we have when we neglect the decaying tail of the internal wave function of the residual nucleus beyond the range of the internal interaction of the residual nucleus. Therefore r_n cannot be smaller than a_0 for $r_p > a_p$. Similar circumstances hold also for the matrices α and β defined by (2.12c) and (2.14d), respectively. For example, the integration with respect to \mathbf{r}_n in the defining integral of $\beta_{p,d}^{(t)}$ must be carried over the interval $r_n > a_0$, with V_p replaced by $V_{p,n}$ if we neglect contributions from $V_{p,\bar{n}}$ again.

Although the direct amplitude $u^{(t)}$ gives us essentially the same result as Butler's for the angular distribution of emerging protons, this amplitude alone does not determine the actual angular distribution. If the matrices α and β , which represent the direct processes for emerging particles due to the overlapping of the internal wave functions in the external region, are small enough, we have for $U_{p\pi I_p m_p}$, besides $u_{p\pi I_p m_p}^{(t)}$, the amplitude due to the formation of the compound nucleus, that is $S(I^{(t)} + I^{(c)})$.

$SI^{(c)}$ is nothing but the reaction amplitude given by the resonance formula of Kapur and Peierls. If the energy of the incident deuteron is high enough so that many levels of the compound nucleus are excited, and also there are many available levels of the residual nucleus into which the compound nucleus can decay, we may use the statistical theory of Hauser and Feshbach⁽¹⁶⁾ and Wolfenstein⁽¹⁷⁾. Then the angular distribution of protons emitted by the compound nucleus tends to be isotropic, because interferences between various levels and those between various partial waves of emerging protons are likely to cancel out to zero, when the average over the energy spread of the incident beam is performed. The interferences between the direct amplitude $u^{(t)}$ and the compound amplitude $u^{(c)}$ also tend to cancel out, because protons with a definite angular momentum can be emitted from many levels of the compound nucleus. In the case of high deuteron energy, therefore, we might say that the peaked angular distribution for small angle of emerging protons is almost determined by the direct amplitude $u^{(t)}$ due to the overlapping

in the external region.

Furthermore, if the incident energy is high enough so that we may possibly use the evaporation model of Weisskopf and Ewing¹⁸, the energy distribution of protons emitted from the compound nucleus is such that it is much larger for low energy side than for high energy side. In this case, therefore, we could say that the angular distribution of emerging protons with their residual nucleus in its low excitation is certainly determined by the direct stripping process.

$SI^{(d)}$ is the reaction amplitude due to the fact that the directly stripped particles interact with their residual nucleus, forming then the compound nucleus. If we take into account only the overlappings between the pairs of the deuteron and target nucleus, the proton and residual nucleus, and the neutron and residual nucleus, we have for $(SI^{(d)})_{p\pi l_p m_p}$

$$(SI^{(d)})_{p\pi l_p m_p} = \sum_{\pi' l' m'} S_{p, \pi'} I_{p'}^{(d)} + \sum_{\nu l_n m_n} S_{p, n} I_n^{(d)} + \sum_{\delta l_d m_d} S_{p, d} I_d^{(d)}. \quad (3 \cdot 11)$$

Of various terms in the right hand side of (3·11), $S_{p, \pi'} I_{p'}^{(d)}$ with $\pi = \pi'$ represents the elastic scattering of the stripped-off protons by the residual nucleus, when summed over l_p, m_p and $l_{p'}, m_{p}'$. When the incident energy is high enough, this elastic scattering of the stripped-off protons becomes essentially the one given by the continuum theory of Feshbach and Weisskopf¹⁹, or the cloudy crystal ball model of Feshbach, Porter and Weisskopf²⁰. When the incident energy of the deuteron is lower and the Q -value for the (d, p) reaction is not very large, this elastic scattering may become the so-called potential scattering, that from a hard sphere, for some partial wave at some energy of the stripped proton. There are also resonance effects of individual levels of the compound nucleus for the stripped protons. These effects on the angular distribution of the pure absorption and the potential or resonance scattering for the stripped protons have been phenomenologically examined by Horowitz and by Tobocman and Kalos¹¹

We must, however, take into account also the contributions from $SI^{(d)}$. For low incident energies, it will show resonance behaviour due to the formation of a definite compound state by the incident deuteron and the subsequent decay into the proton and the residual nucleus. Such process may obscure the effects of the potential or resonance scatterings of the stripped protons by the residual nucleus.

Other terms in the right hand side of (3·11) are the reaction amplitudes for the pair of the proton-its residual nucleus or for the other pairs. We may use the same reasonings as used in the discussion about the effects of $SI^{(c)}$ on the angular distribution. We may say that if the concerning energies are high enough, on the average, these terms do not give much peaked angular distribution. It may be noted, however, that these considerations concern only with the general trends, but not with individual cases, where fluctuations from these general trends may play important roles in determining actual angular distributions.

If the matrices α and β are not negligibly small, they affect the angular distribution in two ways. One is what they have on the compound amplitude $u^{(c)}$, represented by $S \alpha (1 - \beta)^{-1} u^{(d)}$ and $[1 - S \alpha (1 - \beta)^{-1}]^{-1}$. For the effects of $S \alpha (1 - \beta)^{-1} u^{(d)}$, we may

simply rely upon the statistical or continuum theory or the evaporation model, expecting them to be similar to those of $SI^{(d)}$. The effects of $[(1-\beta\alpha(1-\beta)^{-1})^{-1}]$ are expected to make the angular distribution flatter when the concerning energies are high enough, because this factor represents such processes in which outgoing particles emerging from the compound nucleus, on their ways of flying off, give rise to particles incident upon the compound nucleus from other channels over and over again.

The other effect the matrices α and β have on the angular distribution is that represented by the matrix $(1-\beta)^{-1}$. U can be rewritten as

$$U = u^{(d)} + u^{(c)} + (1-\beta)^{-1} \beta (u^{(d)} + u^{(c)}). \quad (3.12)$$

The third term of this expression is the correction due to the fact that outgoing waves of various pairs, on their ways of propagation, excite outgoing waves of other pairs with one another. $(1-\beta)^{-1}\beta$ represents, therefore, the direct processes for emerging particles due to the overlappings in the external region, but with modified amplitudes, namely $(1-\beta)^{-1}\beta$, instead of β itself. This modification of amplitudes for the direct processes for emerging particles arises from the fact that outgoing particles of a pair, say d , are stripped on their ways of flying away, giving rise to emerging particles of another pair, say p , and these emerging particles of the pair p then produce emerging particles of the pair d through pick-up processes. On account of such repeated exchange of amplitudes of outgoing waves, their effective amplitudes for stripping or pick-up processes are modified from β .

When the incident energy is high enough, $u^{(c)}$ has, as mentioned above, appreciable components for the shadow scattering of the incident wave of deuteron, while for other channels it has, on the average, comparatively small components. If the incident energy is low, of various components of $u^{(c)}$, those corresponding to the elastic or nearly elastic scattering of the incident deuteron, including the potential scattering, are likely to dominate. $[(1-\beta)^{-1} \beta u^{(c)}]_{p\pi l_p m_p}$ is, therefore, expected to represent mainly the processes in which deuterons elastically or inelastically scattered by the target nucleus, including the shadow or potential scatterings, are stripped with the modified amplitude.

Of various terms of $[(1-\beta)^{-1} \beta u^{(d)}]_{p\pi l_p m_p}$, $\sum_{\delta l d m_d} [(1-\beta)^{-1} \beta]_{p,d} u_d^{(d)}$ represents the stripping of the deuterons scattered through the direct process. It is rather likely for the incident deuteron to be stripped than to be elastically or inelastically scattered, when the proton or the neutron, but not both, keeping their relative distance larger than the range of their mutual interaction, comes within the range of the interaction with the target nucleus. Thus this sum is expected to represent mainly the effect of the shadow scattering of the incident deuteron due to its stripping. This gives, together with $[(1-\beta)^{-1} \beta u^{(c)}]$ discussed above, some basis for the numerical calculation of Tobocman and Kalos¹¹⁾ with a modified wave method in which the modified wave of the incident deuteron was taken to be the one describing pure absorptions of some partial waves of the incident plane wave of deuteron.

Other terms of $[(1-\beta)^{-1} \beta u^{(d)}]_{p\pi l_p m_p}$ represent elastic or inelastic scatterings of the protons produced through the stripping of the incident deuterons, and also represent direct

(n, p) reactions due to the overlapping in the external region for the neutrons produced through the deuteron stripping, both with the modified amplitudes. These are expected not to play dominant roles in the (d, p) reaction at moderate energies, because the decaying tails of the wave functions of their residual nuclei are not so long as that of the deuteron internal wave function. On the other hand, for (d, p) reactions at very high energies, these may play some important roles. In such a case, however, the notion of the compound nucleus may lose its significance, because waves incident upon the internal region become able to penetrate deeply into the internal region without losing their original modes of motion. Our discrimination between the internal and the external regions, then, may cease to give us useful procedure of approaching the problem, and we might have better to rely upon other methods of approximation.

§ 4. Concluding remarks

We have discussed (d, p) reactions in terms of the resonance theory of nuclear reactions. In order to keep the close relationship between the notion of the compound nucleus and the eigenfunctions of the total Hamiltonian of the system in the internal region, we have treated the effects of the overlappings of the internal wave functions of various reaction products in the external region approximately, as corrections to the ordinary resonance theories. The resultant expression for reaction amplitudes are given in a form such that physical implications involved are clarified as far as possible. It is hoped that this result may help us in interpreting experimental data on nuclear reactions with which loosely bounded particles such as deuterons or residual nuclei with high excitation concern.

Our result is also applicable to other cases besides those mentioned above. As an example, let us take (n, p) reactions. If the energy of the incident neutron is high enough, its wave length becomes of the same order of magnitude as or shorter than the decay distance of the wave function of the proton outside the target nucleus. It then becomes more probable for the incident neutron to excite proton outgoing waves, through the direct interaction with the proton while they are both outside the range of the strong interactions with the other nucleons. This is what we have called the direct process due to the overlapping in the external region. The result of Austern, Butler and McManus⁽²¹⁾ for direct (n, p) reactions with a Born approximation, can thus be interpreted from our formulation as the one due to the direct amplitude arising from the overlapping in the external region. We can also understand naturally from our formulation, the Butler type calculations of the (t, d) stripping or (d, t) pick-up reactions in the current literatures⁽²²⁻²⁴⁾.

In our approximate treatment we have used two steps of approximation. We have neglected the change of the amplitudes of outgoing waves and also neglected the amplitudes of standing waves in the overlapping part of the external region, and we have approximated the boundary conditions imposed on the total wave function of the system on the surface of the internal region. In order to obtain more reliable result, we may use variational principles, with trial functions of the form suggested from the considerations given in this paper.

The author would like to express his sincere thanks to Professors S. Tomonaga and T. Miyajima for their kind interest and encouragement. Thanks are also due to M. Kawai and H. Ui for enlightening and stimulating discussions.

References

- 1) S. T. Butler, Proc. Roy. Soc. (London) **A208** (1951), 559.
- 2) A. B. Bhatia, K. Hung, R. Huby and H. C. Newnes, Phil. Mag. **43** (1952), 485.
- 3) P. B. Daitch and J. A. French, Phys. Rev. **87** (1952), 900
- 4) R. Huby, Proc. Roy. Soc. (London) **A215** (1952), 385
- 5) N. Austern, Phys. Rev. **89** (1953), 318.
- 6) S. Yoshida, Prog. Theo. Phys. **10** (1953), 1.
- 7) Y. Fujimoto, S. Hayakawa and K. Nishijima, Prog. Theor. Phys. **10** (1953), 113.
- 8) J. Horowitz and A. M. L. Messiah, Phys. Rev. **92** (1953), 1326.
- 9) E. Gerjuoy, Phys. Rev. **91** (1953), 645.
- 10) N. C. Francis and K. M. Waston, Phys. Rev. **93** (1954), 313
- 11) W. Tobocman, Phys. Rev. **94** (1954), 1655; W. Tobocman and M. H. Kalos, Phys. Rev. **97** (1955), 132.
- 12) R. G. Thomas, Phys. Rev. **100** (1955), 25.
H. Ui, Prog. Theo. Phys. **16** (1956), 299.
- 13) P. L. Kapur and R. Peierls, Proc. Roy. Soc. (London) **A166** (1938), 277.
- 14) E. P. Wigner and L. Eisenbud, Phys. Rev. **72** (1947), 29; T. Teichmann and E. P. Wigner, Phys. Rev. **87** (1952), 123.
- 15) T. Teichmann and E. P. Wigner, *ibid.*; R. G. Thomas, *ibid.*
- 16) W. Hauser and H. Feshbach, Phys. Rev. **87** (1952), 366.
- 17) L. Wolfenstein, Phys. Rev. **82** (1951), 690.
- 18) V. F. Weisskopf, Phys. Rev. **52** (1937), 295; V. F. Weisskopf and D. H. Ewing, Phys. Rev. **57** (1940), 472, 935.
- 19) H. Feshbach and V. F. Weisskopf, Phys. Rev. **76** (1949), 1550.
- 20) H. Feshbach, C. E. Porter and V. F. Weisskopf, Phys. Rev. **96** (1954), 448.
- 21) A. Austern, S. T. Butler and H. McManus, Phys. Rev. **92** (1953), 350.
- 22) S. T. Butler and E. E. Salpeter, Phys. Rev. **88** (1952), 133.
- 23) H. C. Newnes, Proc. Phys. Soc. (London) **A65** (1952), 916.
- 24) A. Werner, Nuclear Physics **1** (1956), 9.

Theory of Bose-Einstein Condensation of an Imperfect Bose-Einstein Gas*

Takeo MATSUBARA

Research Institute for Fundamental Physics, Kyoto University, Kyoto

Akira MORITA

Physical Institute, Faculty of Science, Tôhoku University, Sendai

Namio HONDA

*Department of Electrical Engineering, Faculty of Engineering,
Tôhoku University, Sendai*

(Received July 30, 1956)

In order to clarify the behavior of the Bose-Einstein condensation, to which various investigations have been reported with results being in conflict with each other, the influence of binary collisions to the Bose-Einstein condensation is treated correctly. As the result, it is shown that the B-E condensation is of the third order, as in the case of an ideal Bose-Einstein gas model.

§ 1. Introduction

Since London's¹⁾ theory of Bose-Einstein (hereafter simply called B-E) condensation, many authors²⁾⁻⁶⁾ have attempted to explain the λ -transition between liquid helium I and liquid helium II in terms of the imperfect B-E gas model.

With the recent progress of observations of physical properties of liquid He³, it seems unquestionable that the λ -transition of liquid He⁴ must be explained as the B-E condensation of a B-E liquid. Discrepancies between the characters of the λ -transition of liquid He⁴ and that of an ideal B-E gas, especially the discrepancy as to the order of the transition, may be explained by properties of an imperfect B-E gas as quantum liquid. But because of the mathematical difficulty of the many body problem with interacting particles, it seems that no one has been able to succeed in perfecting the explanation. Various approximating methods have been proposed in order to avoid this difficulty. Unfortunately their results did not always coincide, depending on the methods used, even though the same order of approximations were adopted.

*) Though an original form of this paper was published by Matsubara in *Busseiron-Kenkyû* (in Japanese) No. 18 (1949), 74 and a revised one was published by Morita and Honda in *Busseiron-Kenkyû* No. 35 (1951), 124, it was approved, after discussions, to be published jointly in the present form.

For example, Schiff⁵⁾ has investigated the influence of binary collisions on the B-E condensation of an imperfect B-E gas. He has shown that, for such an approximation, the mean occupation number n_k for a single atomic energy level is given by

$$\left. \begin{aligned} n_k &= g_k \left\{ \exp \left[\alpha + \beta \epsilon_k - \sum_{k'} n_{k'} G(\mathbf{k}, \mathbf{k}') \right] - 1 \right\}^{-1}, \\ \sum_k n_k &= N \end{aligned} \right\} \quad (1.1)$$

and the mean energy of the system E is given by

$$E = \sum_k \epsilon_k n_k - (1/2) \sum_{k, k'} n_k n_{k'} \partial G(\mathbf{k}, \mathbf{k}') / \partial \beta, \quad (1.2)$$

where $G(\mathbf{k}, \mathbf{k}')$, defined by (2.2) in the next section, is a correction due to binary collision between \mathbf{k} and \mathbf{k}' atoms. In order to avoid the difficulty to solve the non-linear integral equation (1.1), he has assumed that $\alpha - \sum_{k'} n_{k'} G(\mathbf{k}, \mathbf{k}')$ can be approximately replaced by α above the temperature of condensation. As a result, he has shown that the transition is of the first order.

One of the present authors³⁾ (Matsubara) has used the classical approximation

$$\exp(-\beta [H_0 + U]) = \exp(-\beta U) \exp(-\beta H_0) \quad (1.3)$$

in which the quantum mechanical effect due to the non-commutability of the kinetic energy H_0 with the interatomic potential U is neglected, and, furthermore, he has taken into account only the correlation between atoms that are nearest neighbours in a cycle of permutations. Then he has shown that the B-E condensation of an imperfect B-E gas may be of the second order, though some approximations contained in his calculation are open to criticism.

Friedman and Butler¹⁾ have used the identity

$$\exp(-\beta [H_0 + U]) = \lim_{n \rightarrow \infty} \{ \exp(-\beta U/n) \exp(-\beta H_0/n) \}^n \quad (1.4)$$

and have calculated the partition function of liquid He^4 approximately. Thus they have shown that, though the B-E condensation is of the third order, the temperature dependence of the specific heat below the transition point is improved considerably compared with that of an ideal B-E gas model.

Using the method of path integral formulated by himself, Feynman²⁾ has shown that He^4 atoms move very freely among each other with an effective mass m' larger than the true mass of a He^4 atom and, thus, the partition function of liquid He^4 is calculated by

$$Z(\beta) = (K\beta/N!) (m'/2\pi\beta\hbar)^{-3N} \int \rho(\mathbf{X}_1, \dots, \mathbf{X}_N) \sum_{P(N)} \exp[-(m'/2\beta\hbar^2) \sum (\mathbf{X}_i - P(N)\mathbf{X}_i)^2] d\mathbf{X}_1 \dots d\mathbf{X}_N \quad (1.5)$$

with a good approximation. Taking only the correlations between atoms that are nearest neighbours in a cycle of permutations, as one of the present authors (Matsubara) did, he has found that the B-E condensation is of the third order. Here it may be noted that, with exception of the appearance of the effective mass m' , Feynman's partition function (1.5) is substantially equivalent to Matsubara's partition function obtained by using the classical approximation (1.3).

Chester⁶⁾ has proposed an expansion formula of free energy with an interaction force constant and applied it to an imperfect B-E gas. Then, to the first order approximation, he has found that the B-E condensation is of the second order. Furthermore, applying his formula to Feynman's partition function (1.5) he has also obtained the second order. Here it may be noted that the degree of approximation used by Chester is essentially equivalent to that contained in Schiff's calculations taking into account the influence of only binary collisions.

Under these circumstances, therefore, it is very desirable that the correct conclusion is established for each degree of approximation. One already knows that the B-E condensation of an ideal B-E gas is of the third order.¹⁾ In the present paper, we shall give a definite conclusion with respect to the order of the B-E condensation which one can expect when one takes the interatomic force into account only as far as binary collisions

are concerned. (In the following, this approximation is called B-C approximation.) For such an approximation, as Schiff²⁾ has shown, the mean occupation number n_k of each free atomic state k is given by the integral equation (1.1). The difficulty of this problem is that this integral equation is non-linear and that the results do not coincide depending on which approximating method one adopts to solve this integral equation.

In the present paper, an expansion method to solve the integral equation (1.1) is described. Then it is shown that, to the accuracy of B-C approximation, the B-E condensation is undoubtedly of the third order as in the case of an ideal B-E gas.

§ 2. Imperfect B-E gas

In Feynman's partition function of liquid He⁴ (1.5), K_3 is a normalization factor and $\rho(X_1, \dots, X_N)$ represents a density function associated with each configuration and is very similar to the density function $\exp(-\beta U)$ of a classical gas. One may expect the behavior of the B-E condensation almost insensitive to the detailed form of the interaction force. On the other hand, the interatomic force of He atoms which consists of a repulsive core and a weak attractive part, can be substituted by a rigid sphere potential with a fairly good approximation. Therefore we can take the density $\rho(X_1, \dots, X_N)$ to be simply the density function $\exp(-\beta U)$ for a classical gas of impenetrable atoms of diameter a . Thus we see that the approximation

$$\exp(-\beta[H_0 + U]) = \exp(-\beta U) \exp(-\beta m' H_0 / m) \quad (2.1)$$

is very useful for liquid He⁴, where U is the potential energy of a gas composed of impenetrable atoms with diameter a and H_0 is the kinetic energy operator of liquid He⁴.

Now we start with (1.1) and (1.2). In (1.1)*, ϵ_k is the kinetic energy of a free atomic state with wave number vector k and $G(k, k')$ is given by

$$G(k, k') = \exp(\beta[\epsilon_k + \epsilon_{k'}]) V^{-2} \int dX_1 \int dX_2 \{ \exp[-i(kX_1 + k'X_2)] + \exp[-i(kX_2 + k'X_1)] \} \\ \times \{ \exp[-\beta H_{12}] - \exp[-\beta H_{12}^0] \} \exp[i(kX_1 + k'X_2)], \quad (2.2)$$

where H_{12} and H_{12}^0 are the Hamiltonian and the kinetic energy operator of a system of two He⁴ atoms respectively. Then, by considering (2.1), we have

$$\epsilon_k = (\hbar^2 / 2m') k^2$$

and

$$G(k, k') = -V^{-1} \{ Q + J(k - k') \}, \quad (2.3)$$

where

$$Q = \int_0^a 4\pi r^2 dr = (4\pi/3) a^3, \quad (2.4)$$

*) The necessary condition for validity of B-C approximation is, as is shown by Schiff²⁾ that $x = NQ/V < 1$, where N is the total number of atoms, V the volume of the system and Q is given by (2.4). This condition does not hold in real liquid He⁴ where $x \approx 1.7$.

$$J(\mathbf{k}-\mathbf{k}') = \int_{\Omega}^a \exp [i(\mathbf{X}_1 - \mathbf{X}_2) \cdot (\mathbf{k} - \mathbf{k}')] d\mathbf{r}. \quad (2.5)$$

The parameter α is determined by the condition

$$N = \sum_{\mathbf{k}} n_{\mathbf{k}}. \quad (2.6)$$

Now let us put $\rho = N/V$ and

$$V^{-1} \sum_{\mathbf{k}'} n_{\mathbf{k}'} G(\mathbf{k}, \mathbf{k}') = \rho Q + W(\mathbf{k}), \quad (2.7)$$

then we obtain

$$\rho = V^{-1} \sum_{\mathbf{k}} \{ \exp[\alpha + \beta \epsilon_{\mathbf{k}} + \rho Q + W(\mathbf{k})] - 1 \}^{-1} \quad (2.8)$$

and

$$E/V = V^{-1} \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} \{ \exp[\alpha + \beta \epsilon_{\mathbf{k}} + \rho Q + W(\mathbf{k})] - 1 \}^{-1}. \quad (2.9)$$

By considering that the sum $\sum_{\mathbf{k}}'$ over all free atomic states except for the ground state $\mathbf{k}=0$ can be replaced by the integration

$$\{V/(2\pi)^3\} \int \cdots d\mathbf{k},$$

we obtain

$$\rho = V^{-1} \{ \exp[\alpha + \rho Q + W(0)] - 1 \}^{-1} + (2\pi)^{-3} \int \{ \exp[\alpha + \beta \epsilon_{\mathbf{k}} + \rho Q + W(\mathbf{k})] - 1 \}^{-1} d\mathbf{k}, \quad (2.10)$$

$$E/V = (2\pi)^{-3} \int \epsilon_{\mathbf{k}} \{ \exp[\alpha + \beta \epsilon_{\mathbf{k}} + \rho Q + W(\mathbf{k})] - 1 \}^{-1} d\mathbf{k}. \quad (2.11)$$

$W(\mathbf{k})$ is determined by the integral equation

$$W(\mathbf{k}) = V^{-1} \{ \exp[\alpha + \rho Q + W(0)] - 1 \}^{-1} J(0) + (2\pi)^{-3} \int J(\mathbf{k}-\mathbf{k}') \{ \exp[\alpha + \beta \epsilon_{\mathbf{k}'} + \rho Q + W(\mathbf{k}')] - 1 \}^{-1} d\mathbf{k}'. \quad (2.12)$$

§ 3. A solution of the integral equation (2.12)

In this section, we proceed to solve the integral equation (2.12). First we expand the right-hand side of (2.5) as a power series of $\pi = a\mathbf{k}$:

$$J(\mathbf{k}-\mathbf{k}') = A_0 - A_1 |\pi - \pi'|^2 + A_2 |\pi - \pi'|^4 - \cdots, \quad (3.1)$$

where

$$A_0 = Q, A_1 = (1/10)Q, A_2 = (1/280)Q, A_3 = (1/5040)Q, \cdots \quad (3.2)$$

and then the relation

$$A_0 \gg A_1 \gg A_2 \gg A_3 \gg \cdots > 0 \quad (3.3)$$

holds. Next let us assume the expansion

$$\beta^{-1} \{ \rho Q + W(\mathbf{k}) \} = u_0 + (\hbar^2/2m'a^2) [-B_1 \pi^2 + (1-B_1)(B_2 \pi^4 + B_3 \pi^6 + \cdots)] \quad (3.4)$$

and substitute (3.4) into both sides of (2.12). Performing the integration of k' on the right-hand side of (2.12), and equating coefficients of same power terms of k' in both sides of (2.12) we obtain the following simultaneous equations:

$$\left. \begin{aligned} u_0/kT &= A_0 \{V(\exp \alpha^* - 1)\}^{-1} + \rho Q + A_0 I_1 - A_1 I_2 + A_2 I_3 - \dots, \\ (\hbar^2/2m'a^2) (B_1/kT) &= A_1 \{V(\exp \alpha^* - 1)\}^{-1} + A_1 I_1 - (10/3) A_2 I_2 + \dots, \\ (1 - B_1) (\hbar^2/2m'a^2) (B_2/kT) &= A_2 I_2 - 7 A_3 I_3 + (126/5) A_4 I_4 - \dots, \\ (1 - B_1) (\hbar^2/2m'a^2) (B_3/kT) &= \dots, \\ &\dots\dots\dots \end{aligned} \right\} \quad (3.5)$$

where

$$I_\kappa = (2\pi^2)^{-1} a^{-3} \int_0^\infty \pi^{2\kappa} \{\exp[\alpha^* + \theta^{-1} \{\pi^2 + B_2 \pi^4 + B_3 \pi^6 + \dots\}] - 1\}^{-1} d\pi, \quad (3.6)$$

$$\kappa = 1, 2, 3, \dots$$

and

$$\alpha^* = \alpha + u_0/kT, \quad \theta = 2m'a^2 kT / \{\hbar^2(1 - B_1)\}. \quad (3.7)$$

For example,

$$\left. \begin{aligned} I_1 &= \rho - \{V(\exp \alpha^* - 1)\}^{-1} = \left\{ \frac{2\pi m' kT}{\hbar^2(1 - B_1)} \right\}^{3/2} [F_{1/2}(\alpha^*) - (15/4)\theta F_{3/2}(\alpha^*) B_2] \\ &\quad + (105/8)\theta^2 F_{5/2}(\alpha^*) \{(9/4)B_2^2 - B_3\} - \dots, \\ I_2 &= \left\{ \frac{2\pi m' kT}{\hbar^2(1 - B_1)} \right\}^{5/2} \theta [(3/2)F_{3/2}(\alpha^*) - (105/8)F_{5/2}(\alpha^*) B_2] \\ &\quad + (945/16)\theta^2 F_{7/2}(\alpha^*) \{(11/4)B_2^2 - B_3\} - \dots, \\ &\dots\dots\dots \end{aligned} \right\} \quad (3.8)$$

where

$$F_s(\alpha) = \{ \Gamma(s+1) \}^{-1} \int_0^\infty x^s (\exp[\alpha + x] - 1)^{-1} dx = \sum_{n=1}^\infty n^{-(s+1)} \exp(-n\alpha). \quad (3.9)$$

The total energy of the system is given by

$$E/V = (\hbar^2/2m'a^2) I_2. \quad (3.10)$$

Here it may be noted that the expansion method mentioned above is inapplicable when $B_1 \sim 1$, as is seen from (3.7) and (3.8).

§ 4. B-E condensation

One can solve the simultaneous equations (3.5) by successive approximation. When each B_κ is a small quantity, it holds, as is easily seen from (3.3) and (3.5), that

$$B_2/B_1 = O(A_2/A_1), \quad B_3/B_1 = O(A_3/A_1), \dots\dots$$

When $A_\kappa (\kappa \leq 2)$ can be neglected against A_0 and A_1 , one can neglect $B_\kappa (\kappa \geq 2)$ against B_1 . As the first approximation for such a case, we obtain

$$u_0/kT = A_0[V\{\exp \alpha^* - 1\}]^{-1} + A_0 I_1 + \rho Q \quad (4.1)$$

and

$$(b^2/2m'a^2) (B_1/kT) = A_1[V\{\exp \alpha^* - 1\}]^{-1} + A_1 I_1. \quad (4.2)$$

Also we have

$$\rho = [V\{\exp \alpha^* - 1\}]^{-1} + I_1 = [V\{\exp \alpha^* - 1\}]^{-1} + \{2\pi m' kT/[b^2(1-B_1)]\}^{3/2} F_{1/2}(\alpha^*) \quad (4.3)$$

Substituting (4.2) in (4.3), we have

$$\rho = [V\{\exp \alpha^* - 1\}]^{-1} + \{2\pi m' kT/(b^2[1 - (2m'a^2 kT/b^2) A_1 \rho])\}^{3/2} F_{1/2}(\alpha^*). \quad (4.4)$$

(4.4) determines the value of α^* as a function of temperature T for a given value of the density ρ . In the temperature range, in which $\alpha^* > 0$, that is to say, above the condensation temperature, the first term $\rho_0 = [V(\exp \alpha^* - 1)]^{-1}$ of the right-hand side of (4.4) which represents the number density of atoms occupying the ground state is a negligibly small quantity. In such a temperature range, therefore, (4.4) becomes

$$\rho = \{2\pi m' kT/(b^2[1 - (2m'a^2 kT/b^2) A_1 \rho])\}^{3/2} F_{1/2}(\alpha^*). \quad (4.5)$$

Here $F_{1/2}(\alpha^*)$ is a monotonously decreasing function of α^* for non-negative values of α^* and the coefficient of $F_{1/2}(\alpha^*)$ in the right-hand side of (4.6) is a monotonously increasing function of temperature T (as far as $B_1 < 1$). With decreasing temperature under the condition that the density ρ is fixed, therefore, the value of α^* decreases and becomes zero at the temperature T_c determined by

$$\rho = \{2\pi m' kT_c/(b^2[1 - (2m'a^2 kT_c/b^2) A_1 \rho])\}^{3/2} F_{1/2}(0). \quad (4.6)$$

The magnitude of the right-hand side of (4.6) can not be equal to ρ and is always smaller than ρ below this temperature. Thus the remainder of atoms must fall into the ground state. The B-E condensation does occur precisely at the temperature T_c . The density of atoms occupying excited states (i.e., the so-called normal part) below the temperature T_c is given by

$$\rho_n = \{2\pi m' kT/(b^2[1 - (2m'a^2 kT/b^2) A_1 \rho])\}^{3/2} F_{1/2}(0). \quad T < T_c \quad (4.7)$$

Now let us put

$$\left. \begin{aligned} \xi &= \rho_n/\rho, \\ T_0 &= (b^2/2\pi m' k) \{\rho/F_{1/2}(0)\}^{2/3} \end{aligned} \right\} \quad (4.8)$$

and

$$\gamma = (2m'a^2 k/b^2) A_1 \rho, \quad (4.9)$$

then (4.7) and (4.8) are rewritten as

$$T_c/T_0 = 1 - \gamma T_c \quad (4.10)$$

and

$$\xi = [(T/T_0)(1 - \gamma T_c)^{-1}]^{3/2}, \quad T \leq T_c \quad (4.11)$$

respectively.

The detailed behavior of the B-E condensation is determined by the behavior of α^* near T_c . Differentiating (4.6) by temperature T , we obtain

$$\partial \alpha^* / \partial T = -[3 / \{2T(1 - \gamma T)^{2/3}\}] \cdot [F_{1/2}(\alpha^*) / F'_{1/2}(\alpha^*)]. \quad (4.12)$$

Since $F_{1/2}(0) = 2.612$ and $\lim_{\alpha^* \rightarrow 0} F'_{1/2}(\alpha^*) = -F_{-1/2}(0) = -\infty$, we see that

$$\lim_{T \rightarrow T_c} \partial \alpha^* / \partial T = 0. \quad (4.13)$$

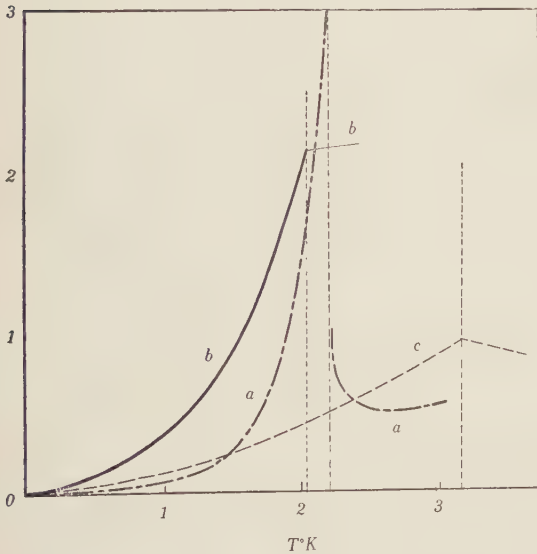


Fig. 1. Specific heat versus temperature curves. Curve *a* shows experimental results for liquid He^4 , *b* is obtained from the present theory, while *c* is obtained from an ideal gas theory.

an ideal B-E gas with the same density as that of real liquid He^4 is 3.14°K . The specific heat versus temperature curve is shown in the Fig. 1 together with those of an ideal B-E gas and of real liquid He^4 .

§ 5. Discussion

In the previous sections, we have shown that the B-C approximation yields the B-E condensation of the third order, as in the case of an ideal B-E gas. This result is contrary to those obtained by Schiff²⁾ and Chester⁴⁾ respectively. The origin of the discrepancy between our calculation and theirs lies in that we take account of the change of the parameter α due to the B-C approximation correctly while they approximate it

This fact means that the B-E condensation mentioned above is of the third order just as that of an ideal gas.

Though the condition of validity of B-C approximation does not hold in real liquid He^4 , it is interesting to compare our results with those of an ideal B-E gas model and observations on real liquid He^4 . If we tentatively put the effective mass m' equal to the true mass of He^4 atom and adopt $a = 2.6\text{\AA}$ (the interatomic potential is zero at $a = 2.6\text{\AA}$ and there is a violent repulsion if the atoms approach more closely than 2.6\AA), we obtain $T_c = 2.02^\circ\text{K}$. On the other hand, the observed λ -temperature is 2.19°K and the condensation temperature of

by the value of the ideal B-E gas. The transition temperature and temperature dependency of the specific heat below the transition point obtained from the present theory are improved considerably compared to an ideal B-E gas model. It must be noted, however, that the condition of the validity of B-C approximation is not satisfied for real liquid He^4 . On the other hand, one may encounter a great amount of mathematical difficulties in extending the theory so as to include higher order collisions.

References

- 1) F. London, *Phys. Rev.* **54** (1938), 947.
- 2) L. I. Schiff, *Phys. Rev.* **59** (1941), 751, 758 and **60** (1941), 362.
- 3) T. Matsubara, *Prog. Theor. Phys.* **6** (1951), 714.
- 4) M. H. Friedman and S. T. Butler, *Phys. Rev.* **98** (1955), 284 and 294.
- 5) R. F. Feynman, *Phys. Rev.* **91** (1953), 1291.
- 6) G. V. Chester, *Phys. Rev.* **94** (1954), 246 and **100** (1955), 455.

Two-Nucleon Problem with Pion Theoretical Potential, I* — Determination of Coupling Constant and Deuteron Problem —

Junji IWADARE,[†] Shoichiro OTSUKI,^{††} Ryoza TAMAGAKI^{††}
and Wataro WATARI^{††}

Department of Physics, Kyoto University, Kyoto^{††}

Department of Physics, Hiroshima University, Hiroshima^{††}

(Received July 7, 1956)

Assuming the pion theoretical potential due to the exchange of one pion in the region outside the pion Compton wave length, the deuteron problem is analysed. Since the deuteron has a large radius on account of its small binding energy, the wave function is strongly affected by the outer part of the potential. This outer part of the pion theoretical potential has the effective coupling constant $g_e^2/4\pi$ between the p -wave pion field and a nucleon as the only variable parameter. Thus we can determine the value of the coupling constant from the deuteron problem as $0.065 < g_e^2/4\pi < 0.090$. This value is in good agreement with that inferred from pion reactions.

The form of the deuteron wave function with the pion theoretical potential and its approximate analytical expression are given.

An information is obtained that the tensor potential must be negative inside the pion Compton wave length in order to give the bound deuteron state, which is consistent with the pion theoretical potential in the same region.

§ 1. Introduction

The second-order nuclear force has recently attracted much more attention of people than in former years. The triplet P -wave phase shift of low energy proton-proton scattering experiments ($E_{lab} < 4.2$ Mev) was found to be negative.¹⁾ This negative phase shift was explained²⁾ as due to the characteristic features of the pion theoretical potential at large distances. The main contributing factor to the P -wave phase shift is the repulsive central part of the second-order nuclear force, which is effective when two nucleons are separated by two or more times the pion range $\hbar/\mu c$ (μ : the pion mass). The pion theory may be expected to be reliable in the region of the outer part of the potential. Therefore the result was regarded as a strong support to the pion theoretical nuclear forces. At the same time, it was found that the quantitative agreement is very good when the coupling constant $g^2/4\pi$ in the pseudoscalar pion theory with P -wave coupling is about $0.06 \sim 0.08$.

Before the above research, many Japanese physicists,²⁾³⁾⁴⁾ had attacked the problem of nuclear forces dividing the configuration space of nucleons into two regions along the line

†) Now at Research Institute for Fundamental Physics, Kyoto University, Kyoto.

*) A part of the content of the present paper is quoted in a preceding letter by the authors with the title "Determination of the Pion Coupling Constant in Nuclear Forces".¹⁴⁾

of the approach proposed by Taketani and others.⁵⁾ In the outer region $r \gtrsim 0.6 \hbar/\mu c$, the interaction was represented by the second-plus-fourth-order potential in the perturbation approximation. On the contrary, in the inner region $r \lesssim 0.6 \hbar/\mu c$, the interaction was represented by arbitrary phenomenological potentials, since the present pion theory is not expected to give any reliable answer when the two nucleons are very close together. By a comprehensive comparison of the theoretical results with experimental data at both low and high energies ($E_{lab} = 100$ Mev), it was found that the pion theory of nuclear forces can reproduce the data *qualitatively* if $g^2/4\pi$ is about 0.08.

After the above research into the 3P -wave phase shift at low energies, it was thought reasonable that in order to give more *quantitative* arguments, one should discriminate the contributions of the second-order potential from those of the fourth- and higher-order ones. The reason was that all methods to construct the potential proposed so far result in the same potential shape as the second-order one at large distances, while each of them gives different shapes of the inner potentials, though the differences are not very large in most states. Thus, the so-called second-order potential seems to lie on a firm ground, while the fourth- and higher-order ones do not. One should discuss about the former as *quantitatively* as possible, while one can do about the latter only *qualitatively*.

Furthermore, it was proved that when the inter-nucleon distance is very large, i.e. under the condition that the two nucleons are nearly free and the momentum transfer is very small, the one-pion-exchange process shown in Fig. 1 gives rise, in the region outside about 1.5 times the pion range, to the same potential shape as the so-called second-order potential, i.e.,

$$V^{(1)} = (g^2/4\pi) (1/3) \mu c^2 (\boldsymbol{\tau}_1 \cdot \boldsymbol{\tau}_2) [(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) + S_{12}(1 + 3/x + 3/x^2)] e^{-x}/x, \text{ for } x \gtrsim 1.5,$$

$$x = \mu cr/\hbar = \kappa r^*, \quad (1)$$

Fig. 1. Diagrams representing (a) the general type of the one-pion-exchange process and (b) the second-order process. A is the complicated vertex part and B and C represent the self-energy parts of the nucleon and pion lines respectively.

where $g^2/4\pi$ is the effective (renormalized) coupling constant of the p -wave pion in the symmetrical pseudoscalar pion theory. We call the potential shape of eq. (1) the *one-pion-exchange potential*.

*) $1/\kappa = 1.415 \times 10^{-13}$ cm corresponding to $\mu = 273 m_e$.

We now divide the configuration space into three regions.* In region I ($x \gtrsim 1.5$), the interaction between two nucleons is represented by the one-pion-exchange potential. Moreover the shape of the one-pion-exchange potential is not altered in the region even if one cuts off the high frequency part of the pion field with a reasonable cut-off momentum $\hbar k_{\text{max}} \geq Mc$ (M : the nucleon mass). In region II ($1.5 \gtrsim x \gtrsim 0.7$), the potential due to two-pion-exchange processes is as large as the one-pion-exchange potential. However, the unambiguous prescription for the calculation of the two-pion-exchange effects is not still established, and various different prescriptions result in different shapes of the potential. Among many possibilities, we have some reasons to consider that the potential constructed by Fukuda, Sawada and Taketani's method⁷ is most reliable. Unfortunately, their potential is sensitive to the cut-off procedure in the region II. The recoil corrections and the S -wave coupling of pions with nucleons also affect the potential in the region II. In region III ($x \lesssim 0.7$), there is little hope to believe that the pion theory at the present stage can give the correct description of the interaction as was pointed out by Taketani and others⁵⁾.

The main purpose of the present work is to determine the value of the coupling constant $g_e^2/4\pi$ of the one-pion-exchange potential. This is indispensable for the quantitative development of the theory of nuclear forces and pion reactions. High energy phenomena would not do much for this purpose, since various effects due to many complicated processes other than the one-pion-exchange are expected to be large. At low energies, though the triplet P -wave phase shift predicts $g_e^2/4\pi$ to be about 0.08, it is not free from rather large uncertainties that come from the unestablished interaction in the regions II and III and from large experimental errors. It is shown in next section how $g_e^2/4\pi$ is determined in the deuteron problem. In § 3, the method of calculation is described. In § 4, the results are discussed. The effective coupling constant is determined as $0.065 < g_e^2/4\pi < 0.090$ from the deuteron problem.

Besides the determination of $g_e^2/4\pi$, the following results are important: The form of the deuteron wave function is almost uniquely determined in the outer region as eq. (10'). The D -state probability of the deuteron when the pion theoretical potential is present in the outer region $x > 1$ is predicted as $0.05 \sim 0.08$, which is larger than the value 0.04 currently adopted. In the inner region $x < 1$ (at least in the region $0.6 < x < 1$), the tensor force is found to have the negative sign in order to make the deuteron state bound, which is consistent with the qualitative properties of the pion theoretical potential.

It deserves a special emphasis that the value of the effective coupling constant determined here in the deuteron problem is in good agreement with the values inferred from pion reactions as will be discussed in § 5.

According to the same line of the approach developed in the present work, the low energy data in the singlet even state will be treated in the forthcoming work II, the p - p scattering at 18 Mev in III and the photodisintegration of the deuteron in IV. The results of all these works together with that of the negative 3P -wave phase shift mentioned

*) For more detailed discussions, see Supplement of Progress of Theoretical Physics, No. 3 (1956).

above,²⁾ some being quantitative and the others qualitative, may constitute the verification of the pion theory of nuclear forces.*

§ 2. Deuteron problem with pion theoretical potential

The one-pion-exchange potential in the triplet even state is

$$V^{(1)} = - (g_e^2/4\pi) \mu c^2 [1 + S_{12}(1 + 3/x + 3/x^2)] e^{-x}/x. \quad (2)$$

It should be noted that the tensor potential is negative and is much stronger than the central one.

As can be seen from Fig. 2, the two-pion-exchange potential that is large in the region II is not yet uniquely determined. Its central part becomes appreciable in the region $x < 1.4$. However, its sign does not coincide if we apply various methods to the derivation. In spite of this difficulty concerning the central part of the two-pion-exchange potential, the deuteron problem, fortunately, can be treated pion theoretically and the effective coupling constant $g_e^2/4\pi$ can also be determined. The reasons will be given in detail a little later.

First, the experimental values of the deuteron parameters are given:

| | |
|--|---|
| the binding energy of the deuteron | $B = 2.226 \pm 0.004 \text{ Mev},$ |
| the triplet effective range | ${}^3r_e = (1.704 \pm 0.030) \times 10^{-13} \text{ cm},$ |
| the electric quadrupole moment of the deuteron | $Q = (2.738 \pm 0.014) \times 10^{-27} \text{ cm}^2,$ |
| the D -state probability of the deuteron ⁸⁾ | $P_D = 0.05 \sim 0.15.$ |

Many works were carried out to estimate mesonic corrections to Q , but unfortunately their results do not agree with one another. Moreover they depend on the behavior of the wave function at small distances, so we can not place too much confidence on their values. Sato and Itabashi⁹⁾ estimated the corrections by the normalized Tamm-Dancoff approxi-

*) For a comprehensive summary, see Supplement of Progress of Theoretical Physics, No. 3 (1956) Part II.

Fig. 2. (See next page as to the figure captions.)

The pion theoretical potential in the triplet even state.

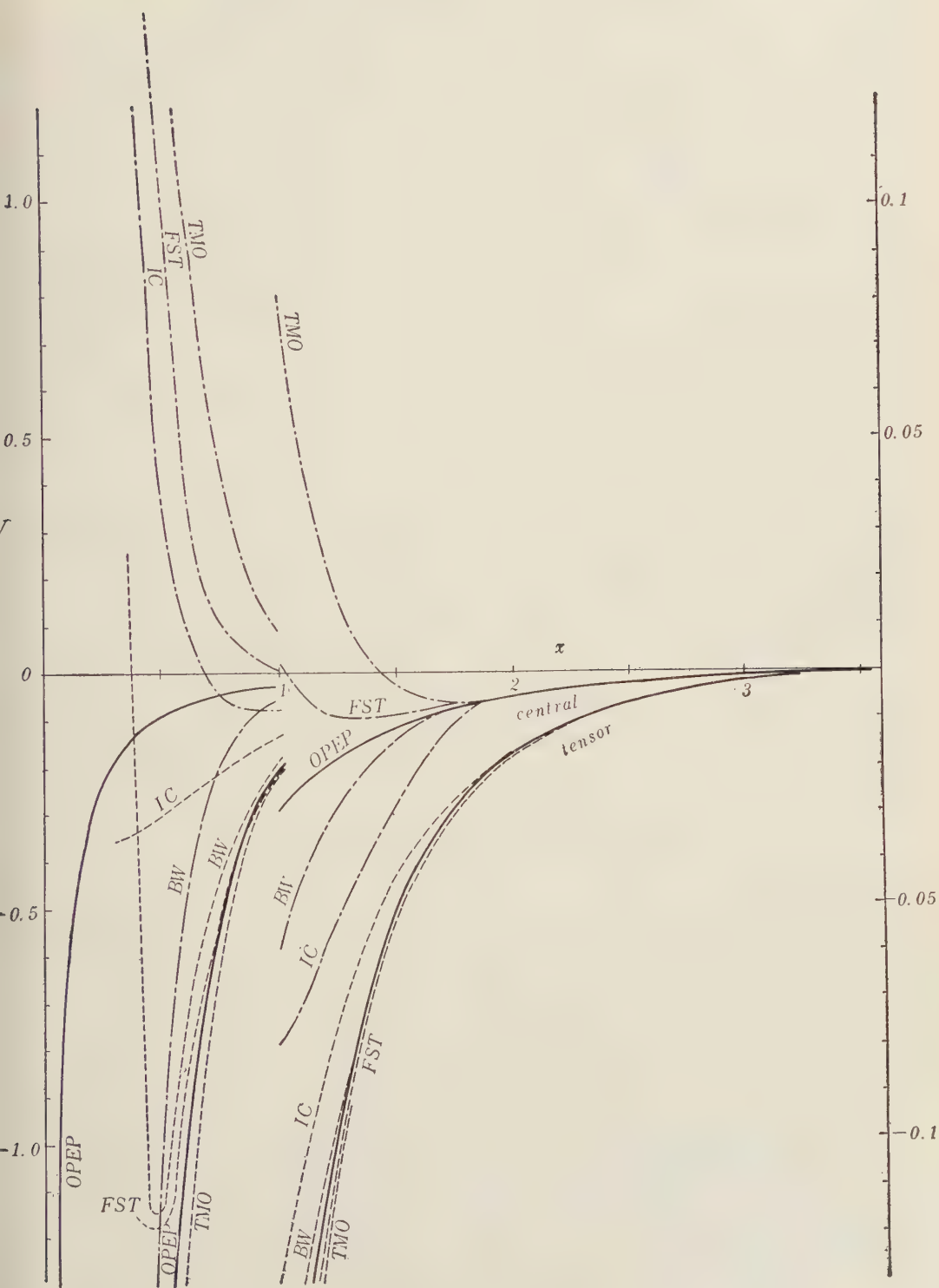
OPEP: The one-pion-exchange potential with $g_e^2/4\pi = 0.08$.

FST: The potential constructed by Fukuda, Sawada and Taketani's method,⁷⁾ with $g_e^2/4\pi = 0.08$. The probability that the nucleons are to be bare is properly taken into account. The high frequency part of the pion field is cut off by the gaussian factor with the cut-off momentum $\hbar k_c = 6\mu c$.

TMO: the $g^2 + g^4$ potential in the perturbation expansion calculated by Take-tani, Machida and Onuma, with $g^2/4\pi = 0.08$.³⁾ The probability part is also expanded in powers of the coupling constant.

BW: The $g^2 + g^4$ potential calculated by Brueckner and Watson with $g^2/4\pi = 0.08$.¹⁰⁾ The probability part is approximated by unity.

IC: The one- and two-pion-exchange potential in the intermediate coupling theory with $g_e^2/4\pi = 0.08$.¹⁰⁾ The cut-off momentum of the gaussian cut-off factor is $\hbar k_c = 4.1\mu c$.



mation. According to their results, the corrections vanish in the static limit and an estimate using the wave function (10') predicts that they are of the order of $4.2 \times 10^{-29} \text{ cm}^2$ i.e., only about 1.5% of the experimental value. Therefore they are very small, and we neglect the mesonic corrections in our discussions.

Usually P_D is estimated from the magnetic moment of the deuteron μ_D , neglecting mesonic corrections. The estimate gives $P_D = 0.039$. However, P_D is very sensitive to the corrections in the magnetic moments, since the value is estimated from the small difference between the large quantities, i. e. $(\mu_p + \mu_n) - \mu_D$, where μ_p and μ_n are the magnetic moments of the proton and the neutron respectively. Machida estimated P_D in another way noticing that the ratio of the hyperfine structure separations of the ground states of deuterium and of hydrogen depends on P_D . The value of P_D above cited is a reasonable guess from his work.⁸⁾

In the presence of the potential $V = V_c + S_{12} V_T$, the Schrödinger equation for the wave function of the deuteron

$$(1/N) (1/x) [u(x) + (1/\sqrt{8}) S_{12} w(x)] {}^3\chi_1$$

(N is the normalization factor of the wave function.)

is given by

$$\begin{aligned} (d^2/dx^2) u(x) + (-\gamma^2 - (M/\mu^2 c^2) V_c(x)) u(x) &= (M/\mu^2 c^2) \sqrt{8} V_T(x) w(x), \\ (d^2/dx^2) w(x) + [-\gamma^2 - (6/x^2) + (M/\mu^2 c^2) (-V_c(x) + 2V_T(x))] w(x) \\ &= (M/\mu^2 c^2) \sqrt{8} V_T(x) u(x), \end{aligned} \quad (3)$$

where $(1/\gamma) = 3.05$ is the deuteron radius corresponding to the binding energy 2.226 Mev. The asymptotic form of the radial wave function $(u(x), w(x))$ outside the range of the forces is

$$\begin{aligned} u(x) \rightarrow u(x) &= A_s e^{-\gamma x}, \\ w(x) \rightarrow w(x) &= A_w e^{-\gamma x} [1 + 3(\gamma x)^{-1} + 3(\gamma x)^{-2}], \end{aligned} \quad (4)$$

where A_s and A_w are constants.

Now we shall explain why the effective coupling constant $g_c^2/4\pi$ can be determined in the deuteron problem.

a) First, we estimate the ratio of the asymptotic amplitudes of $w(x)$ and $u(x)$, i.e., A_w/A_s . The effective range 3r_e is defined by*

$${}^3r_e = (2/\kappa) \int_0^\infty [e^{-2\gamma x} - (u^2(x) + w^2(x))] dx, \quad (5)$$

* Strictly speaking, the effective range 3r_e corresponding to the experimental value is not defined by eq. (5) but by

$${}^3r_e = (2/\kappa) \int_0^\infty [(1 - \frac{x}{{}^3a\kappa}) e^{-\gamma x} - (u_0 u + w_0 w)] dx$$

where ${}^3a = (5.377 \pm 0.023) \times 10^{-13} \text{ cm}$, the triplet scattering length, and (u_0, w_0) is the wave function at zero energy. However, the difference between (5) and (5') is less than a few percent.

where the wave function in this equation is normalized as

$$u^2(x) + w^2(x) \rightarrow \exp(-2\gamma x) \quad \text{as } x \rightarrow \infty.$$

Since $P_D \ll 1$, we can expect $A_D \ll A_S$. Then Eq. (5) becomes

$$\kappa^3 r_e = 2 \int_0^\infty e^{-2\gamma x} dx - \frac{2}{A_S^2 + A_D^2} \int_0^\infty (u^2 + w^2) dx = (1/\gamma) - 2(N^2/A_S^2),$$

where
$$N^2 = \int_0^\infty (u^2 + w^2) dx.$$

Thus we can get from the experimental value of 3r_e as

$$A_S/N \cong \sqrt{2/\{(1/\gamma) - \kappa^3 r_e\}} \cong 1.04.$$

The quadrupole moment Q is defined as

$$Q = (1/\kappa^2 N^2) [(1/\sqrt{50}) \int_0^\infty x^2 u w dx - (1/20) \int_0^\infty x^2 w^2 dx]. \quad (6)$$

Since the deuteron radius is much larger than the pion range, the asymptotic forms $u(x)$ and $w(x)$ approximate $u(x)$ and $w(x)$ well over a wide region of x . As is well known, Q is an "outer" quantity due to the factor x^2 in the integrand of eq. (6), i.e., the main contribution to Q comes from the outer region $x \geq 1$. Thus eq. (6) is approximated by substituting $u(x)$ and $w(x)$ in place of $u(x)$ and $w(x)$. Furthermore its second term may be neglected since $P_D \ll 1$. In this way, we obtain

$$\kappa^2 Q = (1/\sqrt{8}) (A_S A_D / N^2) (1/\gamma^3).$$

From the above equation A_D/N is estimated as

$$A_D/N \cong \sqrt{8} \gamma^3 \kappa^2 Q / (A_S/A_D) = 2\gamma^3 Q \sqrt{(1/\gamma) - \kappa^3 r_e} = 0.013.$$

Therefore, the required ratio A_D/A_S is

$$A_D/A_S \cong 0.01. \quad (7)$$

The approximations made above are rather crude, so this ratio is only qualitative. It is worthwhile to notice the fact the ratio A_D/A_S is a few percent. The more accurate calculation to be performed later will show that $A_D/A_S = 0.025$. However, such an uncertainty in A_D/A_S does not matter to the following arguments.

b) Now we have some knowledge of the form of the two components of the deuteron wave function outside the range of the forces and the ratio A_D/A_S , i.e., Eqs. (4) and (7).

In the region of x where the magnitude of the one-pion-exchange potential becomes larger than the deuteron binding energy or the centrifugal potential, i.e., $x < 2.5$, the two components $u(x)$ and $w(x)$ of the wave function deviate from their asymptotic forms ($u(x)$, $w(x)$). These deviations come mainly from the strong tensor potential as will be explained below.

Let us solve the wave equation (3) starting from outside and proceeding inward.

The deviations of $u(x)$ from $u(x)$ and $w(x)$ from $w(x)$ come from the differences of $(u''(x) - u''(x))$ and $(w''(x) - w''(x))$ respectively. To the former difference $(u''(x) - u''(x))$ contribute two terms, the central contribution $V_C(x)u(x)$ and the tensor contribution $\sqrt{8} V_T(x)w(x)$. The ratio of these two terms is estimated by substituting $u(x)$ and $w(x)$ in place of $u(x)$ and $w(x)$ and using the ratio A_D/A_S of eq. (7). The result is that the ratio $|V_C u|/|\sqrt{8} V_T w|$ is 0.6 at $x=2.0$, 0.3 at $x=1.5$ and 0.1 at $x=1$. Hence we can see that the contribution from the tensor part is larger than that from the central part around these distances. The same situation holds also for $(w''(x) - w''(x))$. The ratio $|V_C w|/|\sqrt{8} V_T u - 2V_T w|$ is 0.02 at $x=2.0$ and 0.04 at $x=1.0$. Consequently the deviations of $u(x)$ from $u(x)$ and $w(x)$ from $w(x)$, or, the behaviors of $u(x)$ and $w(x)$ are determined mainly by the strong tensor force of the one-pion-exchange potential. Needless to say, the potential becomes negligibly small at very large distances, $x \gtrsim 2.5$.

c) As can be seen from Fig 2, the central part of the two-pion-exchange potential becomes quite appreciable in the region $1 \lesssim x \lesssim 1.4$. But the tensor part of the potential is negligibly small in the same region. An estimate similar to b) shows that the effect on $u(x)$ and $w(x)$ due to the central part of the two-pion-exchange potential is less significant than the effect due to the tensor part of the one-pion-exchange potential. Thus we can conclude that $u(x)$ and $w(x)$ in the outer region $x > 1$ are mainly *under the influence of the one-pion-exchange potential, particularly of its tensor potential*. Consequently, the wave function depends critically on the coupling constant $g_e^2/4\pi$ of the one-pion-exchange potential.

d) In the inner region $x < 1$, the potential derived from the pion theory is not reliable quantitatively in the region II ($x \gtrsim 0.7$), nor even qualitatively in the region III ($x \lesssim 0.7$). The potential is expected to affect strongly the behaviors of the inner wave function ($u_{in}(x)$, $w_{in}(x)$). In any case, however, the wave function is compelled to vanish at the origin owing to the boundary condition.

We can assume many forms of the inner wave function ($u_{in}(x)$, $w_{in}(x)$) as long as we know only that they vanish at the origin and pass smoothly at $x=1$ into the outer wave function ($u(x)$, $w(x)$) obtained in a), b) and c). If some of them can reproduce the experimental values of 3r_e , Q and P_D , we regard them as the possible deuteron wave functions yielded by the pion theoretical potential with the assumed value of the effective coupling constant.

e) As the quadrupole moment is an outer quantity, it depends on the detailed behavior of $u(x)$ and $w(x)$ in the outer region $x > 1$, as well as on the amplitudes of the asymptotic forms through $A_S A_D / N^2$. Therefore Q is sensitive to the magnitude of $g_e^2/4\pi$ of the one-pion-exchange potential. On the contrary the detailed behavior of the inner wave function ($u_{in}(x)$, $w_{in}(x)$) hardly affects the magnitude of Q as far as it vanishes at the origin and reproduces the correct value of 3r_e . The effective range 3r_e is an "inner quantity". This gives the measure of the deviation of $u(x)$ and $w(x)$ from the extrapolation of their asymptotic forms $u(x)$ and $w(x)$ as can be seen from the defining equation. (5).

f) The D -state probability P_D depends not only on A_D/A_S and $g_e^2/4\pi$, but also on the

behavior of the inner wave function.

g) If, assuming a definite value of $g_e^2/4\pi$, we find that no inner wave function reproduces all experimental values of 3r_e , Q and P_D , we must be prepared to test another value of $g_e^2/4\pi$,

h) If we find that no values of $g_e^2/4\pi$ fit the experimental data, we might be able to conclude that the symmetrical pseudoscalar pion theory of nuclear forces is incorrect.

However, this is not the case. We can get the wave function of the deuteron reproducing all experimental values of 3r_e , Q and P_D with the pion theoretical potential. At the same time the value of $g_e^2/4\pi$ can be determined.

§ 3. Calculation

We solve the wave equation (3) starting from outside and proceeding inward. There are two linearly independent sets of solutions (u_1, w_1) and (u_2, w_2) which have the asymptotic form (4). For example, we can assume the asymptotic form of (u_1, w_1) such as $(u_1 \rightarrow \exp(-\gamma x), w_1 \rightarrow 0)$ and that of (u_2, w_2) such as $(u_2 \rightarrow 0, w_2 \rightarrow \exp(-\gamma x))$. Then the general solution of eq. (3) that satisfies the boundary condition at large distances (4) is given by a linear combination of the (u_i, w_i) : $(u = u_1 + \alpha u_2, w = w_1 + \alpha w_2)$. The coefficient α in the combination represents the ratio of the asymptotic amplitudes of $w(x)$ to $u(x)$: $\alpha = A_D/A_S$.

We get the above two linearly independent solutions numerically in the outer region $x > 1$ assuming the one-pion-exchange potential in the same region, for three cases of the effective coupling constant $g_e^2/4\pi = 0.06, 0.08$ and 0.10 respectively. The deuteron wave function is given by a linear combination of the two solutions, where the combination coefficient α is expected to be of the order 0.01 as was estimated in § 2, a). The adjustable parameters of the outer wave function are the coupling constant $g_e^2/4\pi$ and the ratio of the asymptotic amplitudes $\alpha = A_D/A_S$.

In the inner region $x < 1$, we do not solve the wave equation for the assumed inner potential but assume the functional form of the wave function itself. The reasons for this description are as follows.

In the inner region around the pion range ($0.7 < x < 1$), the two-pion-exchange potential plays the most important role. However, the signs of the central part of the potential is unfortunately not independent of the method of derivation. For example, by the perturbation approximation (the TMO potential)³⁾, the central potential is strongly positive; Fukuda, Sawada, and Taketani's method gives a weakly positive potential⁷⁾; Brueckner and Watson's method and the intermediate coupling approximation give the negative ones¹⁰⁾ (see Fig. 2). Therefore, it is hard to have the reliable potential shape.

One might think that one could vary the shape and the strength of the inner potential over a wide region and get some useful information. However, when a tensor force is present, the calculation is too laborious and moreover the result would be too complicated to draw easily any inferences regarding pion theoretical nuclear forces.

We shall therefore assume the functional form of the inner wave function $(u_{in}(x),$

$w_{in}(x)$). To do this, we shall attach the conditions to the inner wave function that it vanishes at the origin, continues smoothly to the outer wave function and that it reproduces the effective range 3r_e . Both $u(x)$ and $w(x)$ of the outer wave function with $A_D/A_S=0.01\sim 0.03$ have their maxima near $x=1$ or slightly outer. Therefore, if we once assume the radius of the hard core, that is, if we specify the point x_0 inside which the wave function vanishes, the gross form of the inner wave function can be almost automatically determined.

Once the radius of the hard core x_0 is fixed, a change in the forms of $u_{in}(x)$ and $w_{in}(x)$ produces only a small change in the values of 3r_e and the D -state probability P_D . Also the detailed behavior of $v_{in}(x)$ and $w_{in}(x)$ hardly affects the value of the quadrupole moment Q . Thus we can say that the only important parameter in the inner region in determining the deuteron parameters is the core radius x_0 . Here the core radius x_0 means that inside x_0 there are some interactions that push the wave function outward at such low energies. The interpretation of the interactions at small distances would be outside the potential picture and even outside the applicability of the present theory, since they are expected to be closely connected with the structure of elementary particles.

We calculate Q , 3r_e and P_D using the wave function obtained in the way above mentioned.

$$Q = (1/N^2) (\sqrt{2}/10\kappa^2) \int_0^\infty x^2 (uw - w^2/\sqrt{8}) dx = (Q_{out} + Q_{in}) / (N_{out}^2 + N_{in}^2),$$

$${}^3r_e = \kappa^{-1} \{1/\gamma - 2N^2/(A_S^2 + A_D^2)\} = \kappa^{-1} \{1/\gamma - 2(N_{out}^2 + N_{in}^2)/(A_S^2 + A_D^2)\}, \quad (8)$$

$$P_D = \int_0^\infty w^2 dx / N^2 = \int_0^\infty w^2 dx / (N_{out}^2 + N_{in}^2),$$

where

$$Q_{out} = (\sqrt{2}/10\kappa^2) \int_1^\infty x^2 (uw - w^2/\sqrt{8}) dx, \quad Q_{in} = (\sqrt{2}/10\kappa^2) \int_{x_0}^1 x^2 (uw - w^2/\sqrt{8}) dx,$$

$$N_{out}^2 = \int_1^\infty (u^2 + w^2) dx, \quad N_{in}^2 = \int_{x_0}^1 (u^2 + w^2) dx.$$

The outer quantities Q_{out} and N_{out}^2 depend only on $g_e^2/4\pi$ and A_D/A_S .

Let us consider the dependence of 3r_e and Q on the inner wave function inside $x=1$ when N_{out}^2 and Q_{out} are given. The portion of the inner contribution N_{in}^2 to the norm N^2 is about 20~30%. When N_{in}^2 increases, 3r_e decreases, as is easily seen. When N_{in}^2 increases Q also decreases, since the inner contribution Q_{in} to the numerator of the quadrupole moment is less than 5%. Therefore we can say that the effect of the inner wave function on 3r_e and Q comes mainly through the norm N_{in}^2 and that the larger ratio N_{in}^2/N_{out}^2 gives the smaller 3r_e and Q if the outer wave function is the same.

For the sake of simplicity and generality, we take for $v_{in}(x)$ and $w_{in}(x)$ the quadratic functions in the region $x_0 \leq x \leq 1$ which vanish at a common point x_0 . We calculate Q , 3r_e and P_D according to eq. (8) for $g_e^2/4\pi=0.06$, 0.08 and 0.10 with

$0 \leq x_0 \leq 0.4$ varying the ratio A_D/A_S as a parameter. The results are shown in Fig. 3.

When $g_e^2/4\pi$ decreases the ratio N_{in}^2/N^2 increases since the deviation of $u(x)$ and $w(x)$ in the outer region from the extrapolation of their asymptotic form is reduced owing to the smallness of the potential. Therefore, the smaller value of $g_e^2/4\pi$ gives the smaller values of 3r_e and Q , and vice versa.

As the radius of the hard core increases from $x_0=0$ up to $x_0=0.4$, the inner norm N_{in}^2 is reduced, hence both 3r_e and Q increase by 10~20% and about 7% respectively.

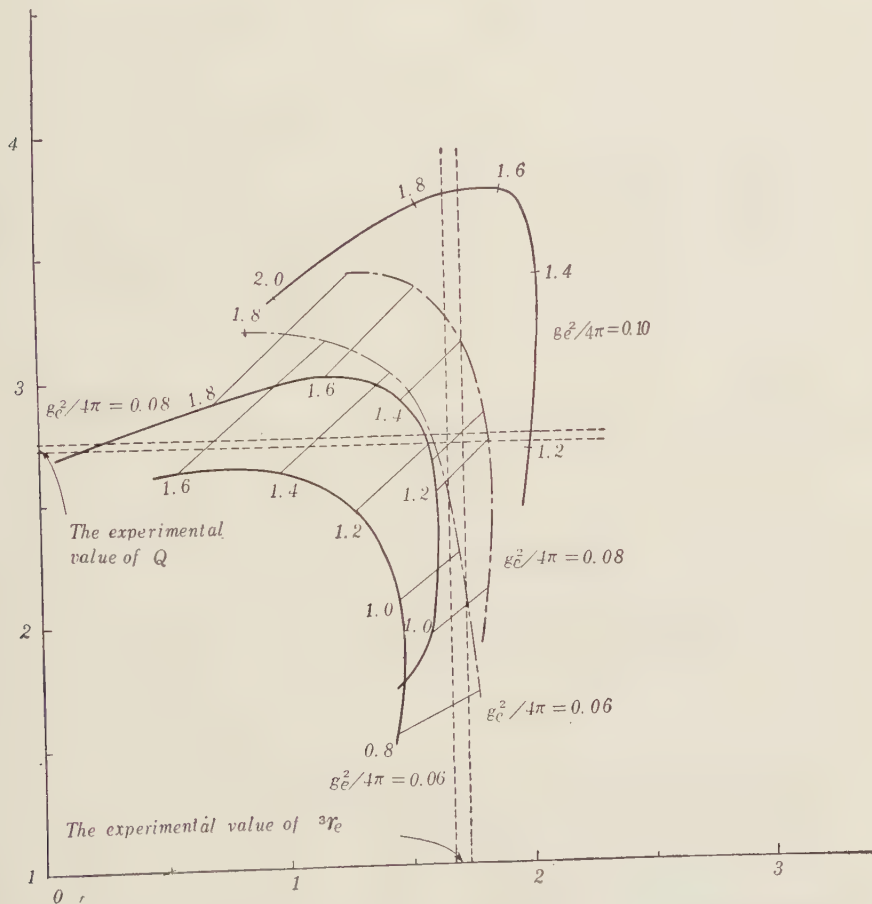


Fig. 3. Q and 3r_e predicted by the deuteron wave function when the one-pion-exchange potential with $g_e^2/4\pi=0.06, 0.08$ and 0.10 respectively are present in the outer region $x>1$.

—: in the absence of a hard-core-like repulsion at small distances.

- - - - -: in the presence of a hard core with the radius 0.4 .

The inner wave function is assumed to be quadratic in x . The ratio of the asymptotic amplitude A_D/A_S is given by the parameter α on the curves by the relation $A_D/A_S=0.01986 \alpha$.

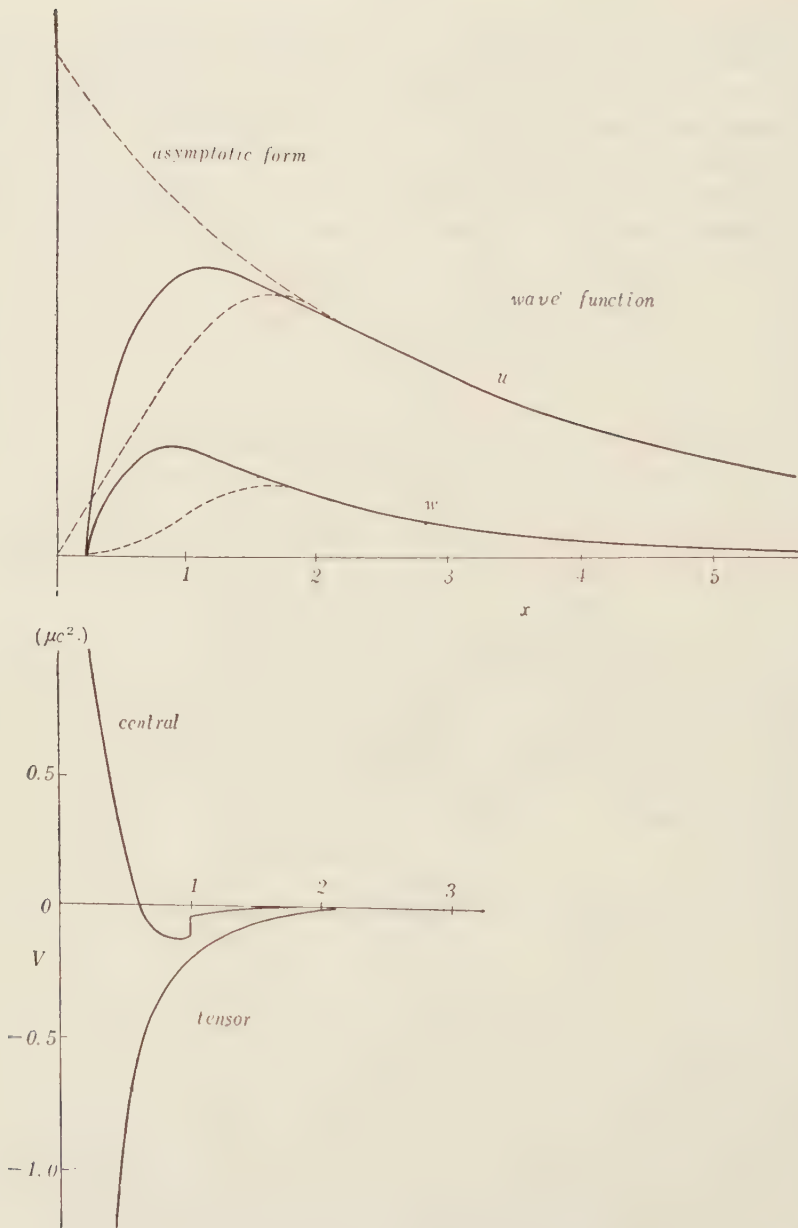


Fig. 4. An example of the deuteron wave function reproducing the correct values of the experimental data when the one-pion-exchange potential with $g_{\pi^2}/4\pi=0.08$ is assumed in the outer region $x>1$. The dotted lines are the wave function obtained by Rarita and Schwinger with a square well potential.¹¹⁾

The potential which influences the wave function is also shown. It is the one-pion-exchange potential in the outer region. In the inner region, the potential is estimated from the wave function.

For the small value of $g_e^2/4\pi < 0.085$, a hard core is required to reproduce 3r_e .

§ 4. Results and discussions

1) Determination of $g_e^2/4\pi$

From Fig. 3 we obtain the upper and lower limits of $g_e^2/4\pi$ outside which it is impossible to reproduce the deuteron data. Namely the value of $g_e^2/4\pi$ is restricted as

$$g_e^2/4\pi = 0.075 \pm 0.010.$$

For larger $g_e^2/4\pi (> 0.085)$, the wave function reproducing the experimental value of 3r_e gives too large Q , e.g., $Q > 3.7 \times 10^{-27} \text{ cm}^2$ for $g_e^2/4\pi = 0.10$, which is about 30% larger than the experimental value. For smaller $g_e^2/4\pi (< 0.065)$, even in the presence of a hard core with the large radius $x_0 = 0.4$, the wave function gives both too small 3r_e and too small Q . It may be possible to obtain a wave function reproducing the correct 3r_e and Q for the values of $g_e^2/4\pi$ which are somewhat smaller than 0.065, if a strong repulsion with the radius larger than 0.4 is assumed. However, it does not seem plausible to assume such a repulsion with the radius as large as one half of the pion Compton wave length. In addition, our argument about the low energy parameters of the singlet S-state in the forthcoming paper II will indicate that $g_e^2/4\pi > 0.07$.

Now we estimate the effects of corrections due to the two-pion-exchange processes to the one-pion-exchange potential in the outer region $x > 1$. The corrections in the outer region are very small for the tensor potential but rather large for the central potential. However, as was stated in § 2, c), the central part of the two-pion-exchange potential in the outer region has little effect on the result.

When we take into account these corrections, the logarithmic derivatives of $u(x)$ and $w(x)$ at $x=1$ change appreciably though the changes in the amplitudes are negligible. Consequently the behavior of the inner wave function is more or less affected by it. If the correction to the central potential is repulsive as in the TMO potential,³⁾ the inner contribution to the norm N_{in}^2 becomes larger, and if it is attractive as in Brueckner and Watson's potential¹⁰⁾ N_{in}^2 becomes smaller. So, in the former case both Q and 3r_e become smaller and in the latter case larger.

Actually, however, even if the central part of the two-pion-exchange potential is as large as in the TMO potential, the one-pion-exchange potential with a slightly larger value of the coupling constant, $g_e^2/4\pi = 0.080 \pm 0.010$, reproduces the deuteron wave function with a good fit to the experimental data. When it is attractive, the fitting value of $g_e^2/4\pi$ becomes somewhat smaller. Fukuda, Sawada and Taketani's treatment predicts the two-pion-exchange potential to be slightly repulsive⁷⁾. It is expected that the slightly repulsive potential has little effect on the value of $g_e^2/4\pi$ determined by assuming the one-pion-exchange potential alone outside $x=1$.

Therefore we can reasonably conclude that the value of the effective coupling constant lies in the following region,

$$0.065 < g_e^2/4\pi < 0.090. \quad (9)$$

Besides the determination of $g_e^2/4\pi$, some important results about the deuteron wave function, P_D , etc., are obtained based on the fact that the one-pion-exchange potential is the main part of the potential in the outer region $x > 1$.

2) Deuteron wave function

The deuteron wave function in the ground state obtained along the way mentioned in § 2 is illustrated in Fig. 4 for the case when $g_e^2/4\pi = 0.08$. This wave function reproduces the experimental data as

$$\left. \begin{aligned} g_e^2/4\pi &= 0.080, \\ A_D/A_S &= 0.0248, \\ x_0 &\cong 0.2, \end{aligned} \right\} \rightarrow \left\{ \begin{aligned} Q &= 2.74 \times 10^{-27} \text{ cm}^2, \\ r_e &= 1.70 \times 10^{-13} \text{ cm}, \\ P_D &= 0.067. \end{aligned} \right. \quad (10)$$

When the wave function is normalized as $1 = \int_0^\infty (u^2 + w^2) dx$, it is approximated by the analytical expression:

$$\begin{aligned} u &= 1.039 \exp(-0.328x) - 1.392 \exp(-2.360x), \\ w &= 0.02624 \{1 + 3/(0.328x) + 3/(0.328x)^2\} \exp(-0.328x) \\ &\quad - (1.298/x^2) \exp(-0.962x), \end{aligned} \quad (10')$$

where the expression for $w(x)$ is not valid for $1 \lesssim x$. This approximate expression may be useful to calculate various deuteron reactions.

For the sake of comparison, the wave functions obtained by Rarita and Schwinger¹¹⁾ with a square well potential of the radius 2.80×10^{-13} cm (nearly equal to $2 \times (b/\mu c)$) and depths of $V_C = -13.9$ Mev and $V_T = 10.8$ Mev are also shown in Fig. 4. Two remarkable features are to be noticed. One is that the asymptotic ratios A_D/A_S coincide with each other in both cases. This is natural because both wave functions can reproduce the quadrupole moment. The other is that both $u(x)$ and $w(x)$ with the pion theoretical potential have their peaks inner than those with the square well potential. This is a consequence of the fact that the pion theoretical potential is not a "short-tailed" potential as the square well but a "long-tailed" one. From Fig. 4, we can expect that the deuteron wave function has its S - and D -peaks near or slightly outside the pion range.

The ratio of the asymptotic amplitude A_D/A_S in the presence of the one-pion-exchange potential is almost completely determined by the outer quantity Q as

$$A_D/A_S = 0.0245 (1 \pm 0.03). \quad (11)$$

3) D -state probability, P_D

From Fig. 5 we can see that the D -state probability P_D with the pion theoretical potential is restricted as

$$0.05 < P_D < 0.08. \quad (12)$$

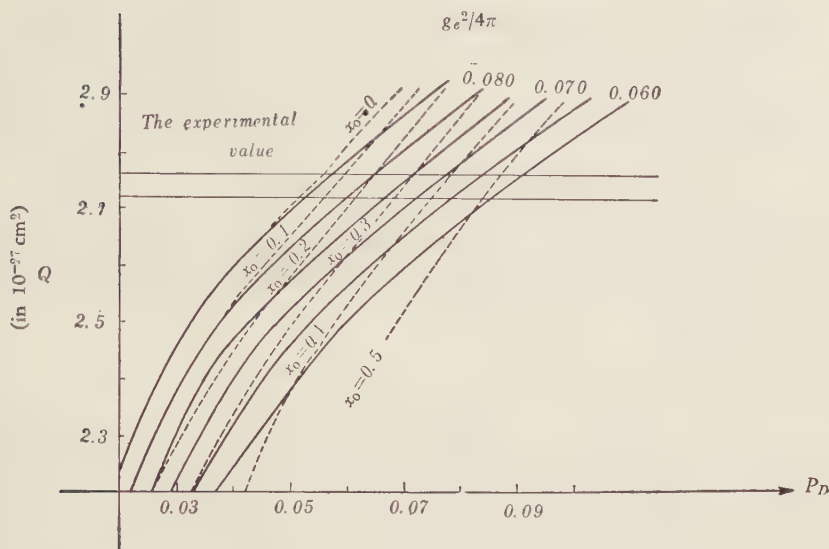


Fig. 5. Full curves represent the relation between Q and P_D , where 3r_e is fitted to the experiment. Dotted curves are equi-hard-core-radius curves. Notice that $P_D > 0.05$.

Rarita and Schwinger's wave function shown in Fig. 4 gives $P_D = 0.04$. The small deviation of $w(x)$ (using the pion theoretical potential) from the extrapolation of the asymptotic form, and hence its steep rise are due to the long-tailed character and the weak central force of the one-pion-exchange potential. If the central force is strongly attractive, the term $V_C(x)w(x)$ in the Schrödinger equation (3) tends to reduce $w(x)$ and consequently reduces P_D . Therefore the large P_D is a characteristic result of the existence of the one-pion-exchange potential in the outer region $x > 1$. The large P_D is consistent with the result obtained by Machida⁹⁾ as discussed in § 2.

4) Potential inside pion range

In order to have some idea on the interaction in the inner region $x < 1$, we investigated what kind of potential is required to give the wave function shown in Fig. 4. This can be done by solving the algebraic equation (3) regarding V_C and V_T as unknown variables. The result is also shown in Fig. 4.

The tensor potential in the inner region is negative and strong. The central potential is weakly attractive near the pion range and becomes highly repulsive inside. A comprehensive investigation shows that the latter result concerning the central potential is not completely general and may fail for certain form of inner wave functions. The result depends on the value of $g_e^2/4\pi$ through the logarithmic derivatives and the amplitudes of $u(x)$ and $w(x)$ at $x=1$, on the effect of the two-pion-exchange potential in the outer region, and on the assumed form of the inner wave function. On the contrary the former result concerning the tensor potential is valid for the case $0.065 < g_e^2/4\pi < 0.090$:

Only a strongly negative tensor potential in the inner region $x < 1$ can surpass the centrifugal potential and make the inner D -wave component $w_{in}(x)$ vanish at the origin. The strongly negative tensor potential in the region II is compatible with the field theoretical result concerning the two-pion-exchange potential as can be seen from Fig. 2.

§ 5. Concluding remarks

We have determined the effective pion-nucleon coupling constant $g_e^2/4\pi$, from the deuteron problem by placing due reliance on the one-pion-exchange potential outside the pion range.

The results is that

$$0.065 < g_e^2/4\pi < 0.090.$$

It is to be noticed that the deuteron problem developed above distinctly determines the upper limit of $g_e^2/4\pi$. Unless various corrections to the quadrupole moment are quite appreciable, which is hardly supposed, the upper limit will not be changed. On the contrary, the lower limit is not so certain. If there is a strong repulsion with the radius larger than one half of the pion range, the smaller value of $g_e^2/4\pi$ is preferable. However, our analysis of the low energy parameters of the singlet even state in the forthcoming work II will indicate that $g_e^2/4\pi$ is larger than at least 0.07.

Thus we can say that if and only if the effective coupling constant of the one-pion-exchange potential is

$$g_e^2/4\pi = 0.080 \pm 0.010, \quad (13)$$

all low energy phenomena of the two-nucleon system, i.e., the deuteron parameters, the singlet even state data and the negative triplet P -wave phase shift,²⁾ can be explained without any exception.*

The value of the effective coupling constant can also be inferred independently from the experiments on pion reactions. Bernardini and Goldwasser analysed the photo-pion production experiment at the threshold energy using the Kroll-Ruderman's theorem. After taking into account the Coulomb interaction, they obtained the result that¹²⁾

$$g_e^2/4\pi = 0.073 \pm 0.007. \quad (14)$$

On the other hand, the plot of the phase shift of the dominant P -wave with the angular momentum $J=3/2$ and the isotopic spin $I=3/2$ in the pion-nucleon scattering at low energies against energy shows also that¹³⁾

$$g_e^2/4\pi \sim 0.08. \quad (15)$$

The values of the pion-nucleon coupling constant obtained from the pion reactions (14) and (15) agree well with (13) determined in the nuclear force problems. This

*) More comprehensive discussions of the low energy phenomena from the standpoint of the pion theory may be found in Supplement of Progress of Theoretical Physics, No. 3 (1956), Part II.

agreement may constitute one of the conceivable successes of the P -wave pion theory of nuclear forces and pion reactions.

Beside the determination of $g_\pi^2/4\pi$, the following results are important: The form of the deuteron wave function is almost uniquely determined in the outer region $x > 1$ as is shown in Fig. 4 and by eq. (10'). The D -state probability P_D is given as $0.05 < P_D < 0.88$ when the one-pion-exchange potential is present in the outer region. In the inner region $x < 1$ the tensor potential is negative and strong which is consistent with the qualitative prediction of the pion theoretical potential.

References

- 1) H. R. Worthington, J. N. McGruer and D. E. Findley, *Phys. Rev.* **90** (1953), 899.
- 2) S. Otsuki and R. Tamagaki, *Prog. Theor. Phys.* **12** (1954), 806; **14** (1955), 52.
- 3) M. Taketani, S. Machida and S. Onuma, *Prog. Theor. Phys.* **7** (1952), 45.
- 4) S. Fujii, J. Iwadare, S. Otsuki, M. Taketani, S. Tani and W. Watari, *Prog. Theor. Phys.* **11** (1954), 11.
S. Otsuki and S. Fujii, *Prog. Theor. Phys.* **12** (1954), 521.
- 5) M. Taketani, S. Nakamura and M. Sasaki, *Prog. Theor. Phys.* **6** (1951), 581.
- 6) K. Hiida, J. Iwadare and S. Machida, *Prog. Theor. Phys.* **15** (1966), 189.
- 7) N. Fukuda, K. Sawada and M. Taketani, *Prog. Theor. Phys.* **12** (1954), 156.
K. Inoue, S. Machida, M. Taketani and T. Toyoda, *Prog. Theor. Phys.* **15** (1956), 122.
- 8) S. Machida, *Prog. Theor. Phys.* **9** (1953), 683.
- 9) I. Sato and K. Itabashi, *Prog. Theor. Phys.* **13** (1955), 341; K. Itabashi, Lecture at the annual meeting of the Physical Society of Japan held at Sendai, July 1956.
- 10) K. A. Brueckner and K. M. Watson, *Phys. Rev.* **92** (1953), 1023.
Y. Nogami and H. Hasegawa, *Prog. Theor. Phys.* **15** (1956), 137.
- 11) W. Rarita and J. Schwinger, *Phys. Rev.* **59** (1941), 436.
- 12) G. Bernardini and E. L. Goldwasser, *Phys. Rev.* **95** (1955), 857. See also reference 13.
- 13) G. F. Chew and F. E. Low, *Phys. Rev.* **101** (1956), 1570, 1579.
- 14) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Prog. Theor. Phys.* **15** (1956), 86.

Two-Nucleon Problem with Pion Theoretical Potential, II — Singlet Even State* —

Junji IWADARE,[†] Shoichiro OTSUKI,^{††} Ryoza TAMAGAKI^{††}
and Wataro WATARI^{††}

Department of Physics, Kyoto University, Kyoto^{†,††}

Department of Physics, Hiroshima University, Hiroshima^{††}

(Received July 7, 1956)

Low energy parameters in the singlet even state are analysed from the point of view of the pion theory. It is shown that the properties which are required for the potential in the singlet even state are as follows: The one-pion-exchange potential in the outer region, a strong attractive force in the region from the pion range to one half of it, which is consistent with the prediction by the two-pion-exchange potential, and a strong repulsion at small distances at least up to $x=0.20$. The value of the effective coupling constant of the one-pion-exchange potential is limited as $g_e^2/4\pi > 0.07$ by the present analysis. This result, together with that of the article I, gives an allowed region for the coupling constant as $0.07 < g_e^2/4\pi < 0.09$.

§ 1. Introduction

In the article I under the same title¹⁾, we treated the deuteron problem from the stand-point of the pion theory, and at the same time, we determined the value of the effective coupling constant of the one-pion-exchange potential as $0.065 < g_e^2/4\pi < 0.090$. The purpose of the present work is to investigate various consequences of the pion theory of nuclear forces in the singlet even state, adopting the established one-pion-exchange potential in the outer region.

It is found that the experimental data at low energies in the singlet even state can be reproduced if the potential has the following properties: The one-pion-exchange potential with $g_e^2/4\pi \sim 0.08$ in the outer region, a strong attractive potential in the region from the pion range to one half of it ($0.5 \lesssim x \lesssim 1.0$), which is consistent with the prediction by the two-pion-exchange potential, and a strong repulsion at small inter-nucleon distances at least up to $x=0.2$, that pushes the wave function outside. At the same time it can be concluded that the effective coupling constant $g_e^2/4\pi$ of the one-pion-exchange potential is larger than 0.07.

The method of calculation is mentioned in § 2. In § 3, the result is discussed. In § 4, some remarks are given.

[†] Now at Research Institute for Fundamental Physics, Kyoto University, Kyoto.

* Some part of the result of the present work has already been reported in this journal.¹⁾

§ 2. Experimental data and the method of calculation

In the shape independent approximation, the p - p scattering experiment gives the effective range as²⁾

$$^1r_e(pp) = (2.65 \pm 0.07) \times 10^{-13} \text{ cm.} \quad (1)$$

The n - p scattering experiment gives the scattering length as³⁾

$$^1a(np) = (-23.69 \pm 0.06) \times 10^{-13} \text{ cm,} \quad (2)$$

while for the effective range, there is a rather large uncertainty, i.e.,

$$^1r_e(np) = (2.1 \sim 2.8) \times 10^{-13} \text{ cm,} \quad (3)$$

although the uncertainty was fairly reduced by recent experiments.¹⁾ The scattering length in the \bar{p} - p system $^1a(pp)$ which is the counterpart of $^1a(np)$ can not be directly given by the experimental data. It can be estimated by assuming the shape of the nuclear potential, and consequently depends on the assumed shape of the potential. Although we do not know the exact shape of the potential, it may be a reasonable guess from Table I that $^1a(pp)$ lies somewhere in the region

$$-20 \times 10^{-13} \text{ cm} < ^1a(pp) < -15 \times 10^{-13} \text{ cm} \quad (4)$$

for a reasonable shape of the potential.

Table I. The p - p scattering length $^1a(pp)$ and the effective range $^1r_e(pp)$ for various potential shapes, r_0 being the core radius. (All figures in the Table are in the unit of 10^{-13} cm.)

Square well

| r_0 | 0 ^{5a)} | 0 ^{5b)} | 0.3 ^{5b)} | 0.48 ^{5b)} | 0.65 ^{5b)} |
|-------------|------------------|------------------|--------------------|---------------------|---------------------|
| $^1r_e(pp)$ | 2.65 | 2.61 | 2.65 | 2.65 | 2.65 |
| $^1a(pp)$ | -17.7 | -17.1 | -15.8 | -15.6 | -15.3 |

Hulthén

| r_0 | 0 ^{5a)} | 0 ^{5c)} | 0.3 ^{5c)} | 0.6 ^{5c)} | 1.2 ^{5c)} |
|-------------|------------------|------------------|--------------------|--------------------|--------------------|
| $^1r_e(pp)$ | 2.65 | 2.74 | 2.65 | 2.61 | 2.60 |
| $^1a(pp)$ | -19.5 | -18.9 | -17.1 | -16.4 | -15.8 |

In order to make a detailed discussion, we also have to take into account the possible effect of the shape parameter P in the effective range theory. We do not know, however, the precise value of the shape parameter of the potential, since we do not know the exact shape of it, in particular, in the singlet even state. To make the matters worse, the shape parameter depends on the behavior of wave function at small distances which is strongly affected by the unknown interaction in the region III. Therefore, we guess that the magnitude of the shape parameter of the actual potential is of the same order as those of the usual phenomenological potentials. Thus, from Table IX of Jackson and Blatt's work²⁾ and Table I, we can expect that

$$^1r_e(pp) = (2.50 \sim 2.85) \times 10^{-13} \text{ cm.} \quad (1')$$

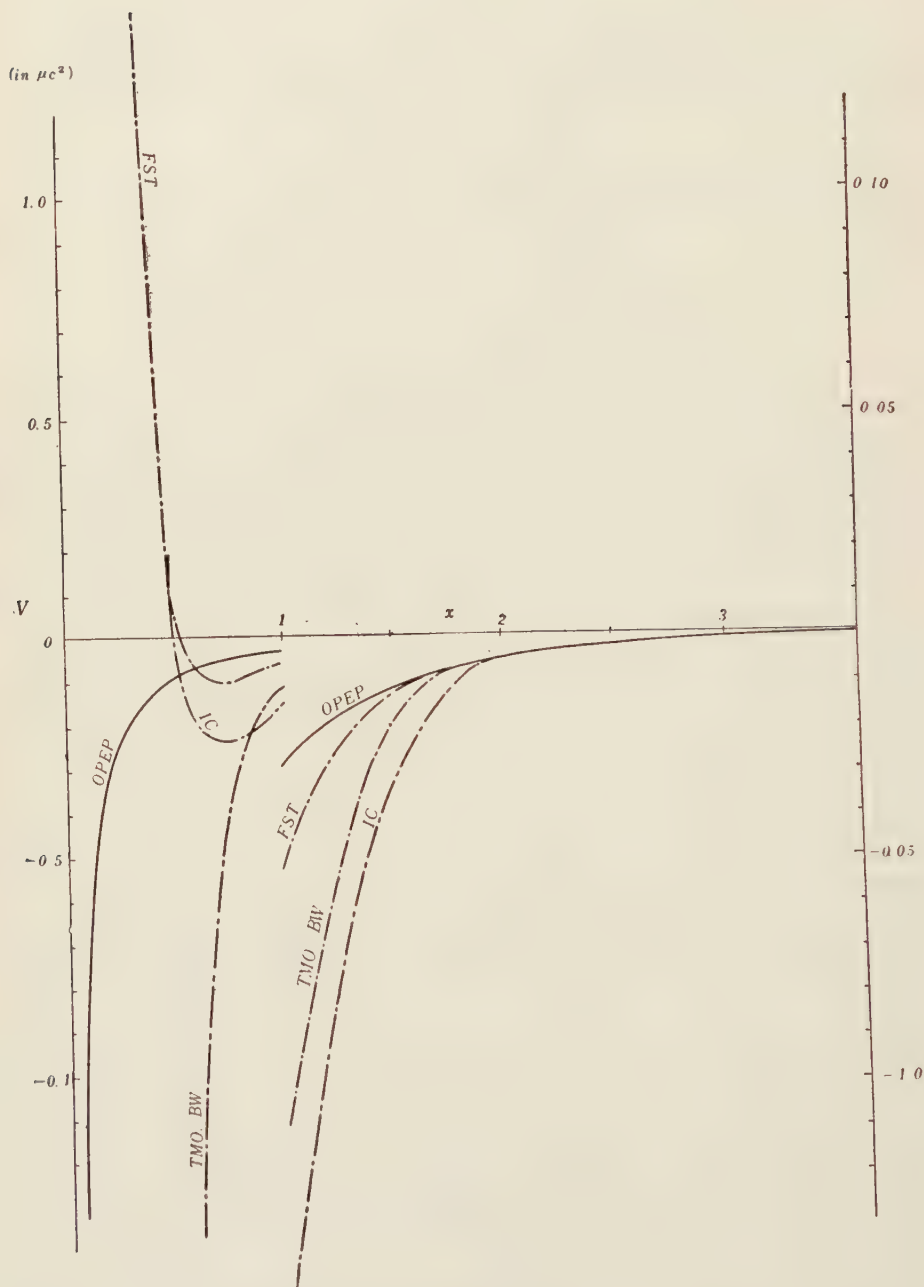


Fig. 1. The pion theoretical potential in the singlet even state.

OPEP: The one-pion-exchange potential with $g_\pi^2/4\pi=0.08$.

FST: The potential constructed by Fukuda, Sawada and Taketani's method, with $g_\pi^2/4\pi=0.08, 6\alpha^{(6b)}$. The probability that the nucleons are to be bare is properly taken into account. The high frequency part of the pion field is cut off by the gaussian factor with the cut-off momentum

Note: Ordinates -1.0 and -0.1 should be exchanged respectively.

Assuming the charge independence of nuclear forces, we use the experimental data of the p - p system throughout the present work, i.e., (4) and (1'). One might think that the apparently large difference between $^1a(np)$ and $^1a(pp)$ would invalidate the charge independence hypothesis. But in fact this difference has only a small effect on our result since the fact that both $|^1r_e(pp)/^1a(pp)|$ and $|^1r_e(np)/^1a(np)| \ll 1$ implies that the potentials in question have a virtual level only slightly above zero energy.

The analysis developed below is an application of the effective range theory to the pion theoretical potential. As is well known, in the effective range theory the potential has two adjustable parameters to reproduce the low energy data of the scattering length 1a and the effective range 1r_e . Since the outer part of the pion theoretical potential or the one-pion-exchange potential has already been established quantitatively, one has to find the two adjustable parameters in the inner potential. Unfortunately, the shape of the two-pion-exchange potential depends seriously on the cut-off factor of the high frequency part of the pion field as can be seen from Fig. 1. Therefore, in the present analysis we adopt the shape of the square well inside $x=1.0$ and regard its depth V_0 as one of the adjustable parameters. The other adjustable parameter is the radius x_0 of a hard core which we assume at small inter-nucleon distances. The core radius is a very useful parameter of the inner potential since a small change of the core radius causes a large change in the inner wave function. The square well radius $x=1$ is chosen merely for convenience' sake. The effect due to the two-pion-exchange potential which is expected to be appreciable even outside $x=1$ will be discussed later.

Thus we assume the potential in the singlet even state as

$$\begin{aligned}
 V &= -(g_e^2/4\pi) \cdot \mu c^2 e^{-x}/x \quad (\text{the one-pion-exchange potential}) \\
 &\qquad\qquad\qquad \text{for } x > 1, \\
 &= V_0 \quad (\text{a constant}) \qquad\qquad\qquad \text{for } 1 > x > x_0, \\
 &= 0 \qquad\qquad\qquad \text{for } x_0 > x,
 \end{aligned} \tag{5}$$

where V_0 and x_0 are the adjustable parameters.

Then, we solve the Schrödinger equation numerically with the one-pion-exchange potential from outside proceeding inward up to $x=1$. We attach the boundary condition to the wave function so that it reproduces the scattering length correctly. The effective coupling constant is taken as $g_e^2/4\pi = 0.065 \sim 0.090$ in accord with the result in I¹⁾. At $x=1$ the wave function is smoothly continued to that inside. The first node point of the wave function gives the core radius x_0 .

$\hbar k_c = 6\mu c$.

TMO: $g^2 + g^4$ potential in the perturbation expansion calculated by Taketani, Machida and Onuma, with $g^2/4\pi = 0.08^{(6c)}$. The probability part is also expanded in powers of the coupling constant.

BW: The $g^2 + g^4$ potential calculated by Brueckner and Watson with $g^2/4\pi = 0.08^{(6a)}$. The probability part is approximated by unity. The difference between the TMO and the BW potentials are negligibly small in the singlet even state. The difference of the FST result from these two is due to the cut-off of the high frequency part.

IC: The one- and two-pion-exchange potential in the intermediate coupling theory with $g_e^2/4\pi = 0.08^{(6e)}$. The cut-off momentum of the gaussian cut-off factor is $\hbar k_c = 4.1\mu c$.

§ 3. Results and discussions

The result of the calculation of § 2 is shown in Fig. 2.

We can see from Fig. 2 that in the inner region there must be a very strongly attractive potential of the order of 100 Mev to reproduce the experimental value of the effective range. The magnitude of the depth V_0 is larger by a factor of at least 10 than that of the one-pion-exchange potential at $\kappa=1$. The fact that the presence of the strong attraction in the inner region is a necessary consequence of the rather large pion range ($\sim 1.4 \times 10^{-13}$ cm) can be explained as below.

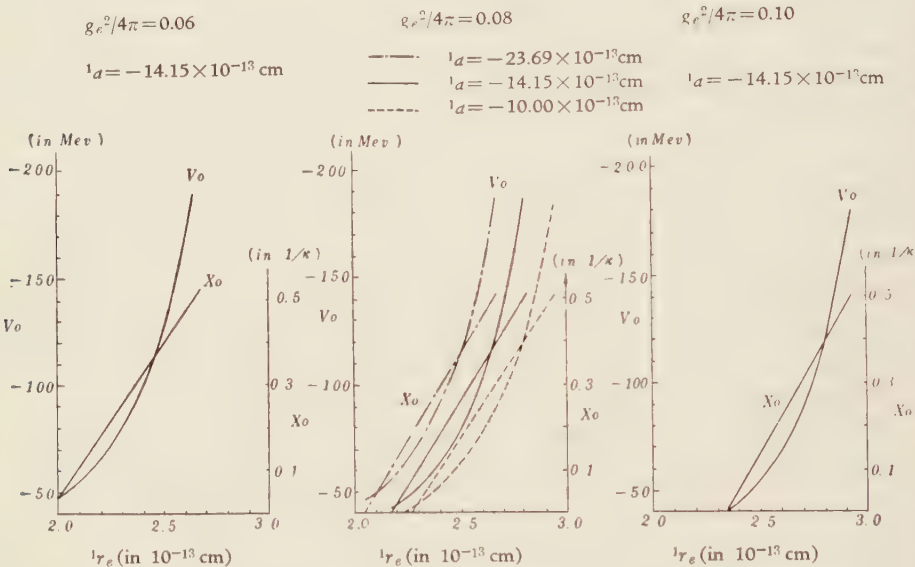


Fig. 2. The depth of the inner potential V_0 and the core radius x_0 of eq (5) are plotted against r_e . For the case of $g_e^2/4\pi=0.08$, three values of 1a are considered. It is expected from the experimental data that

$$\begin{aligned} -20 < ^1a < -15 \times 10^{-13} \text{ cm} \\ 2.5 < r_e < 2.85 \times 10^{-13} \text{ cm.} \end{aligned}$$

Let us suppose that the one-pion-exchange potential is valid up to the origin—a very fictitious assumption only for illustration. Then we have the conventional Yukawa poten-

Table II. Parameters of the Yukawa potential with the pion range that gives a fit to the observed scattering length

| | | |
|--|------------------------------------|------------------------------------|
| $V = -(g_e^2/4\pi) \cdot \mu c^2 e^{-\kappa} / \kappa$ | | |
| 1a | $-23.7 \times 10^{-13} \text{ cm}$ | $-15.0 \times 10^{-13} \text{ cm}$ |
| s | 0.914 | 0.869 |
| $g_e^2/4\pi$ | 0.202 | 0.192 |
| r_e | $3.88 \times 10^{-13} \text{ cm.}$ | $4.14 \times 10^{-13} \text{ cm.}$ |

tial with the pion range. In order to give a fit to the observed scattering length, or to reproduce the correct asymptotic wave function at zero energy, the depth of the Yukawa potential is uniquely determined. The depth parameter s defined in Blatt and Jackson's work⁷⁾ and the corresponding $g_\pi^2/4\pi$, which is about twice larger than the current value, are listed in Table II. As is shown in the same table, the values of 1r_e calculated with the Yukawa potential are also much larger than the experimental one.

We can draw two interesting conclusions from the result in Table II: The actual

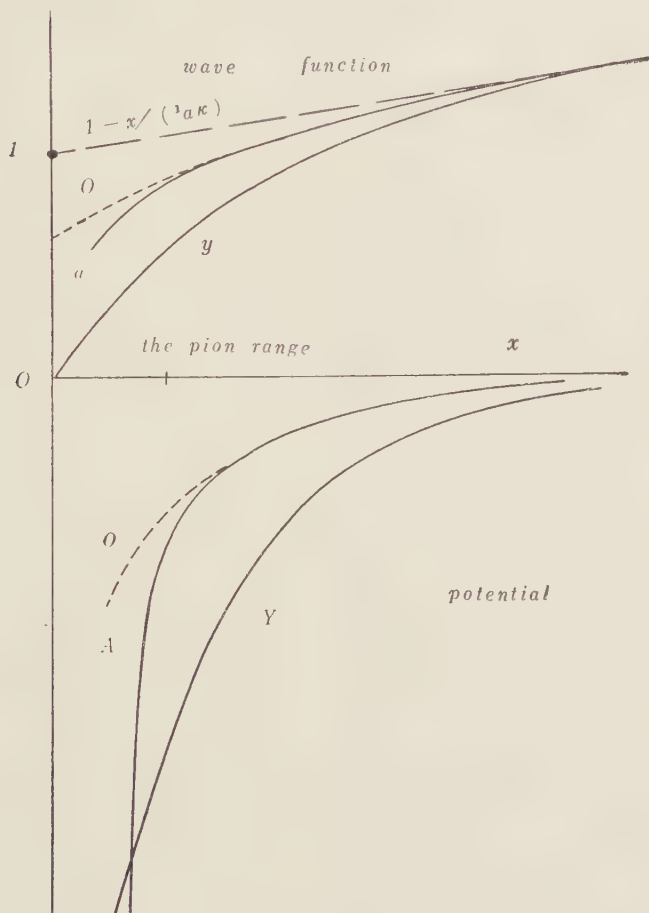


Fig. 3. Schematic wave functions in the singlet even state at zero energy, which reproduce the scattering length.

Y : The Yukawa potential with the pion range in Table II with $g_\pi^2/4\pi=0.2$.

y : Wave function corresponding to the potential Y . The wave function gives too large 1r_e .

O : The one-pion-exchange potential with $g_\pi^2/4\pi=0.08$.

o : Wave function corresponding to the potential O , which does not vanish at the origin.

A : The actual potential including the two-pion-exchange potential.

a : Wave function corresponding to the potential A .

The actual potential is weaker in the outer region and stronger in the inner region than the Yukawa potential Y .

Note: O in the upper region should be read as small letter "o"

O at the origin should be read as zero "0"

potential which should have the pion range in the outer region and reproduce the effective range correctly, is (1) weaker in the outer region and (2) stronger in the inner region than the Yukawa potential in Table II.

These conclusions can be reached as follows. Fig. 3 shows the wave functions schematically when they are under the influence of the Yukawa potential in Table II and that of the actual potential respectively. Since the observed value of 1r_e is much smaller than those given by the Yukawa potentials, the behavior of the actual wave function must be such that its deviation from the extrapolation of the asymptotic form $1-x/(ak)$ is less than the deviation in the case of the Yukawa potential. Therefore the actual potential must be weaker in the outer region. This is the conclusion (1). Then this fact makes it necessary that the actual wave function decreases more rapidly at small distances in order that it may vanish at the origin. Therefore the actual potential must be more attractive than the Yukawa potential in the inner region. This is the conclusion (2).

Here we shall return to the discussions on the result in Fig. 2. We remember that the result was obtained by assuming that the one-pion-exchange potential is valid in all the region $x > 1$. Therefore, we have to estimate the effect in the outer region $x > 1$ due to the two-pion-exchange potential. The estimate is important because the two-pion-exchange potential is as large as the one-pion-exchange potential even in the region $1 < x < 1.5$, as can be seen in Fig. 1. The result of the estimate using the perturbation theoretical potential^(6c) is that when we assume the $g^2 + g^1$ potential in the outer region $x > 1$, the depth of the inner potential V_0 is reduced by a factor smaller than 2 in the case of $g^2/4\pi = 0.06$ while it is reduced by a factor of several in the case of $g^2/4\pi = 0.10$. In both cases, however, V_0 is much deeper than the one-pion-exchange potential. Thus, the conclusion that there must be a strong attractive potential in the inner region is not modified by the inclusion of the two-pion-exchange potential.

Here, it is worth while noticing that if the attraction in the inner region is effective only near the origin, it may be unable to reproduce the sufficient magnitude of the effective range. In the presence of the weak one-pion-exchange potential ($g_e^2/4\pi \sim 0.08$), the wave function at zero energy deviates not so much from asymptotic form in the outer region. (Fig. 3.) Therefore the contribution to the integral 1r_e is expected to be small in the outer region. Indeed, the contribution to the effective range from the outer region $x > 1$ is estimated as only 0.5×10^{-13} cm (about one fifth of the experimental value) even for $g_e^2/4\pi = 0.10$ and for $^1a = -14 \times 10^{-13}$ cm. This is of course an over-estimate, because a larger absolute value of 1a makes the contribution smaller for the same outer potential and a smaller value of $g_e^2/4\pi$ also makes the contribution smaller for the same 1a . The two-pion-exchange potential in the outer region $x > 1$ gives a little contribution to 1r_e , even if one assumes that it is as strong as that derived by the perturbation method without cut-off. The two-pion-exchange potential gives an appreciable change in the logarithmic derivative of the wave function near the pion range, but not in the value of the wave function itself. Thus, the contribution to 1r_e is expected to come mainly from the inner region. Further, if the attractive force in the inner region is effective only near the origin, it may be unable to reproduce the sufficient magnitude of the effective range. Therefore,

it is expected that the inner attractive force may be spread up to the pion range.

Next, we can infer from Fig. 2 that there exists a hard core of the radius of at least 0.25 at small inter-nucleon distances. With a weak outer potential such as the one-pion-exchange potential with $g_e^2/4\pi \sim 0.08$, we can not reproduce the experimental value of the effective range without a hard-core-like repulsion. If any repulsive force is not assumed, the largest value of 1r_e is estimated as at most 2.4×10^{-13} cm with $g_e^2/4\pi = 0.10$ and ${}^1a = -14 \times 10^{-13}$ cm, i.e., in the optimum case. Only repulsion at small distances pushes the wave function outward and makes the inner contribution to 1r_e large. The inclusion of the g^4 potential in the outer region reduces the smallest core radius to about 0.2.

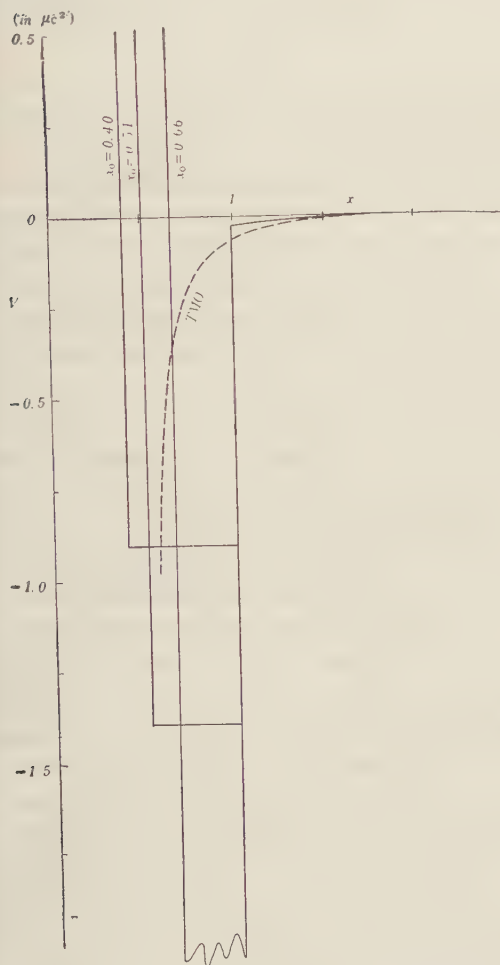


Fig. 4. The pion theoretical potential, Eq. (5), for $g_e^2/4\pi = 0.06$, which gives ${}^1a = -14.15 \times 10^{-13}$ cm, as shown in Fig. 2.

Finally, we can find a lower limit of the effective coupling constant $g_e^2/4\pi$ of the one-pion-exchange potential. The shape of the potential in (5) in the case of $g_e^2/4\pi = 0.06$ is plotted in Fig. 4. One will find a hard core with a radius as large as one half of the pion Compton wave length and a very deep well between the hard core radius and the pion range. However, such a too deep attractive potential near the pion range is hardly accepted from the pion theoretical point of view. It is at least ten times deeper than the $g^2 + g^4$ potential in the perturbation approximation. Even if the effect of the g^4 potential in the outer region is taken into account, the depth V_0 is reduced at most to a half of its original value in the case of $g_e^2/4\pi = 0.06$.

If $g_e^2/4\pi \sim 0.07$, then the depth V_0 is estimated larger by a factor of about 3 than that of the $g^2 + g^4$ potential with $g_e^2/4\pi \sim 0.07$ at $x=1$, if the effect of the g^4 potential in the outer region is taken into account. A larger value of the coupling constant $g_e^2/4\pi > 0.07$ seems natural, since, in such a case, the depth V_0 and that of the pion theoretical potential at $x=1$ have the same order of magnitude.

In Eq. (5), we assumed the square well as the inner potential for simplicity.

However, even if we assume a more plausible shape for the potential, i.e., an increasing function of x , all of our conclusions are not changed, since any inner potential of such a shape gives smaller 1r_c than the square well for the same 1a , $g_c^2/4\pi$ and the core radius x_0 .

§ 4. Concluding remarks

We can summarise the properties of the potential in the singlet even state as follows. The one-pion-exchange potential with the effective coupling constant $g_c^2/4\pi = 0.080 \pm 0.010$ in the outer region, a strong attraction in the region from the pion range to a half of it which is consistent with the prediction by the two-pion-exchange potential, and a strong repulsion at small distances at least up to about $x=0.2$ that pushes the wave function outside.

It must be pointed out that the strong repulsion does not necessarily mean a hard core but means only some interaction that pushes the wave function outward. For example, we may represent it by a condition on the logarithmic derivative of the wave function at some small distance without adopting the potential picture as Eq. (5) in such an inner region. Here lies an essential difference between our stand-point and Lévy's one in his proposal of the "hard core potential"²⁸⁾. Lévy considered that there were some indications of repulsive interaction at small distances for all states. This repulsive terms came from the so-called contact terms, which are spin and τ -spin independent, and have a range of the order \hbar/Mc . He retained the qualitative features of these interaction terms, which were represented by a hard core of a radius of $(\mu/M)^{1/2}(\hbar/\mu c)$. However, we can not agree with Lévy for his interpretation. Contrary to him, the hard core that we have introduced in the present work has nothing to do with today's pion theoretical consequences. No theory at the present stage can predict any definite results on the interactions at small distances and we should not give undue reliance on the field theoretical result.

In 1951, Jastrow introduced a hard core in the singlet even state⁹⁾ in order to give a fit to the p - p scattering angular distribution and to reduce the calculated n - p total cross section at high energies of a few hundred Mev. After his work it came into fashion to introduce a hard core into the problems of nuclear forces. Now it must be emphasized that the introduction of the hard core in the present work is based on the low energy experimental data and on the reliable consequences of the pion theory at large inter-nucleon distances. Though Jastrow's idea is interesting, one should not hastily connect our hard core with Jastrow's one that is based on the analysis of high energy data, since at high energies new aspects of the pion-nucleon interaction, the structure of elementary particles and other complicated effects would appear.

Combining with the result obtained in I, we can conclude that the effective coupling constant of the one-pion-exchange potential is

$$g_c^2/4\pi = 0.080 \pm 0.010. \quad (6)$$

The one-pion-exchange potential with this value of $g_c^2/4\pi$ predicts a rather small singlet D -wave phase shift below several tens Mev. The small singlet D -wave phase shift is very

favorable to explain the isotropic angular distribution of the p - p scattering¹⁰⁾.

In the present work we could not make any quantitative discussion. One of the important reasons for this fact is that the two-pion-exchange potential has not been completely determined yet. Unfortunately, this two-pion-exchange potential is expected to be large even outside the pion range. If the two-pion-exchange potential were exactly known, an analysis along the line of the present work would give much more information on nuclear forces. Another reason that prevents quantitative discussions is the rather large uncertainty in the experimental data to be compared with the theoretical result. A part of this uncertainty comes from the unknown shape of the pion theoretical potential, the dependence on which appears in the effective range theory. It is highly desirable to re-determine the low energy parameters of the singlet even state making full use of our knowledge of the singlet even state potential discussed in the present paper.

References

- 1) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Prog. Theor. Phys.* **16** (1956), 455.
- 2) J. D. Jackson and J. M. Blatt, *Rev. Mod. Phys.* **22** (1950), 77.
- 3) M. T. Burgy, G. R. Ringo and D. J. Hughes, *Phys. Rev.* **84** (1951), 1160.
- 4) R. E. Fields, R. L. Becker and K. A. Adair, *Phys. Rev.* **94** (1954), 389.
C. L. Storrs and D. H. Frisch, *Phys. Rev.* **95** (1954), 1252.
- 5a) J. Schwinger, *Phys. Rev.* **78** (1950), 135.
- 5b) S. Otsuki, *Soryusiron-Kenkyu* (mimeographed circular in Japanese) **5** (1953), 777.
- 5c) E. E. Salpeter, *Phys. Rev.* **93** (1953), 994.
- 6a) N. Fukuda, K. Sawada and M. Taketani, *Prog. Theor. Phys.* **12** (1954), 156.
- 6b) K. Inoue, S. Machida, M. Taketani and T. Toyoda, *Prog. Theor. Phys.* **15** (1956), 122.
- 6c) M. Taketani, S. Machida and S. Onuma, *Prog. Theor. Phys.* **7** (1952), 45.
- 6d) K. A. Brueckner and K. M. Watson, *Phys. Rev.* **92** (1953), 1023.
- 6e) Y. Nogami, and H. Hasegawa, *Prog. Theor. Phys.* **15** (1956), 137.
- 7) J. M. Blatt and J. D. Jackson, *Phys. Rev.* **76** (1946), 18.
- 8) M. M. Lévy, *Phys. Rev.* **88** (1952), 725.
- 9) R. Jastrow, *Phys. Rev.* **81** (1951), 165.
- 10) S. Otsuki and S. Fujii, *Prog. Theor. Phys.* **12** (1954), 521.
- 11) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Prog. Theor. Phys.* **15** (1956), 86.

The Creation of an Electron Pair by a Fast Charged Particle

Toshiyuki MUROTA*, Akira UEDA* and Hajime TANAKA**

**Department of Physics, Kyoto University, Kyoto*

***Department of Physics, Hokkaido University, Sapporo*

(Received July 9, 1956)

Using the Feynman-Dyson method, the cross section for the creation of an electron pair by a fast charged particle is calculated but in a manner more precise than that of Bhabha and others. The differential cross section obtained here is valid as long as the energies of participant particles are large compared with the respective rest masses. The ambiguities in Bhabha's calculation are also examined. It is concluded that the theoretical value must be compared with the experimental results produced by high energy electrons ($\gtrsim 10$ Bev).

§ 1. Introduction

The creation of an electron pair by a charged particle of spin $1/2$ was investigated by Bhabha¹⁾, Nishina et al. and others²⁾, and their results have been applied for analysing highly energetic electromagnetic phenomena in cosmic rays underground. The results of such analyses seem to show that quantum electrodynamics is valid even at extremely high energy, say, 10^{15} eV, though the comparison of the theories with the experiments is done indirectly³⁾.

However, recent developments of the experimental techniques of photographic emulsion have made it possible to measure directly the cross section of the process in question, called trident. Koshiha and Kaplon^{*1)} have indicated that the experimental value of the cross section for tridents produced by a high energy electron is in disagreement with the theoretical value given by Bhabha.

It is an interesting and important problem whether there *really* exists a discrepancy between the experimental and the theoretical results, because it is believed that quantum electrodynamics gives a correct description of electromagnetic phenomena extensively. In order to clarify this point it will be necessary on the one hand that the experimental analyses of tridents be performed more accurately while on the other hand that the theoretical results be derived as strictly as possible following quantum electrodynamics.

As is well known, Bhabha's results have been extensively used for analyzing cosmic ray phenomena. In his approach the incident particle is regarded as a classical one moving along a straight line with uniform velocity, the field of which is replaced by a classical field. This approximation is essentially based on the same assumption as used in

* They cited the similar experimental results obtained by Freier and Naugle (unpublished).

the Williams-Weizsäcker method and gives a correct description only when the energy transferred from the incident particle to the created electron pair is small compared with the energy of the incident particle. On the other hand, the process called "first order process" by Bhabha, i.e., the diagrams D' in the figures, is neglected in B. It is clear that this neglect may be allowed when the mass of the incident particle is large compared with that of an electron (see § 4), but is not justified when the incident particle is in fact an electron. For the latter case the estimation of the contribution from the diagrams D' and also the exchange effect are not treated quantitatively. Therefore it seems necessary to recalculate the cross section of this process more strictly over the whole range of the transferred energy by using a quantum electrodynamical treatment and estimating quantitatively the errors arising from the rather rough treatment in B.

In this paper we shall calculate the process in question using the Feynman-Dyson method, whereby the incident charged particle will be treated quantum dynamically but the target charged particle will be regarded as a fixed Coulomb field. In the following section we shall derive a general formula for the transition probability, explain the treatment used in performing our calculations, and then derive the differential cross section for the process in question. We shall derive the total cross sections in § 3; in § 4 we shall discuss the effects which must be taken into account when the incident particle is an electron, i.e., the exchange effect and the contributions from the diagrams D' which are neglected in § 2. In the final section we shall discuss our results and compare them with the experiments.

§ 2. Differential cross section

(a) General formula for the cross section

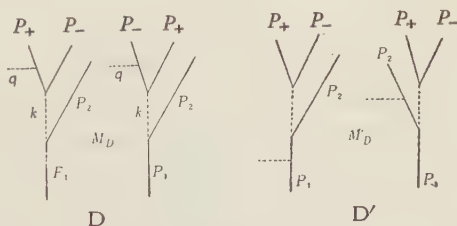
Throughout this paper we use the natural unit, $\hbar=c=1$, and the following notations: $P_1(\mathbf{p}_1, iE_1)$, $P_2(\mathbf{p}_2, iE_2)$: The initial and the final energy-momentum four-vectors of the incident charged particle.

$P_{\pm}(\mathbf{p}_{\pm}, i\epsilon_{\pm})$: The energy-momentum four-vectors of the positron and the electron, respectively.

$k(\mathbf{k}, i\epsilon) = P_1 - P_2$: The energy-momentum four-vector of the virtual photon.
 m and μ : The rest mass of the electron and the incident particle, respectively.

The Feynman-Dyson diagrams for the present process in the lowest order consist of four diagrams as given in the figures. However, we take only the diagrams D. The reason for neglecting the matrix element $M_{D'}$ from the diagrams D' will be made clear in § 4.

From the diagrams D, the matrix element M_D is expressed as



$$\begin{aligned}
M_D = & -Ze^{16\pi^2} \frac{i}{q^2 k^2} u(P_2) \gamma_\mu u(P_1) \bar{u}(P_-) \left[\gamma_\mu \frac{i(\gamma, P_- - k) - m}{D_-} (e\gamma) \right. \\
& \left. + (e\gamma) \frac{i(\gamma, k - P_-) - m}{D_+} \gamma_\mu \right] u(-P_-) \cdot 2\pi \delta_0(P_1 - P_2 - P_+ - P_-) \\
\equiv & K \cdot 2\pi \delta_0(P_1 - P_2 - P_+ - P_-),
\end{aligned} \quad (1)$$

where

$$\begin{aligned}
q &= p_1 - p_2 - p_+ - p_-, \\
D_\pm &= (k - P_\pm)^2 + m^2,
\end{aligned} \quad (2)$$

and $\delta_0(P_1 - P_2 - P_+ - P_-)$ is the fourth component of the four dimensional delta function which expresses the conservation of energy.

The transition probability per unit time, w , is thus

$$w = 2\pi \sum |K|^2 \delta_0(P_1 - P_2 - P_+ - P_-) \rho_F, \quad (3)$$

where \sum represents the summation over the spin directions of the final state and the average over the spin directions in the initial state, and ρ_F is the density of the final state for the three particles, i.e.,

$$\rho_F = (m^2/\epsilon_+ \epsilon_-) d\mathbf{p}_+ d\mathbf{p}_- d\mathbf{p}_2 / (2\pi)^9, \quad (4)$$

where spinors of the incident charged particle are normalized as $u^* u = 1$, while those of the electron pair as $\bar{u} u = 1$.

Dividing the expression (3) by the velocity of the incident particle $|\mathbf{p}_1|/E_1$, we obtain from (3) and (4) the differential cross section,

$$d\sigma = 2\pi \frac{E_1}{|\mathbf{p}_1|} \sum |K|^2 \delta_0(P_1 - P_2 - P_+ - P_-) \frac{m^2}{\epsilon_+ \epsilon_-} \frac{d\mathbf{p}_+ d\mathbf{p}_- d\mathbf{p}_2}{(2\pi)^9}. \quad (5)$$

(b) Calculation of $\sum |K|^2$

In order to calculate $\sum |K|^2$, we introduce a new coordinate system in which the z -axis is parallel to the direction of the propagation of the virtual photon, but not to that of the incident particle, i.e.,

$$z\text{-axis} \parallel \mathbf{k} = \mathbf{p}_1 - \mathbf{p}_2. \quad (6)$$

Henceforth we shall call this coordinate system S-system. The advantage of introducing the S-system is not only that the virtual photon can be naturally separated into the transverse photon and the longitudinal and the scalar ones so that the physical meaning of the calculating procedure is made much clearer, but also that the angular integrations in the final state are more precisely carried out than was done in B. By virtue of this procedure we shall be able to obtain a differential cross section which is valid as long as the energies of participant particles are large compared with the respective rest masses.

In the following discussion we use two approximations: (i) the energies of all the particles are relativistic, i.e., large compared with their rest masses; (ii) the small angle

approximation, since by virtue of the denominators q^2 , k^2 , D_{\pm} , contained in K the main contributions to the cross section come from regions where the angle of each particle is very small. Thus, for example,

$$|\mathbf{p}_1| \sin \theta_1 \cong |\mathbf{p}_1| \theta_1 \equiv p_{1\perp}, \quad |\mathbf{p}_1| \cos \theta_1 - E_1 \cong -(\mu^2 + p_{1\perp}^2)/2E_1. \quad (7)$$

Here the magnitude of the transverse component of each momentum is, as is easily seen, of the same order as that of each rest mass (see (12)).

K consists of two factors, one of which is related with the incident charged particle and the other with the electron pair, i.e.,

$$K = -Ze^4 16\pi^2 A_\mu B_\mu, \quad A_\mu = i\bar{u}(P_2)\gamma_\mu u(P_1), \quad (8)$$

$$B_\mu = \frac{1}{q^2 k^2} \bar{u}(P_-) \left[\gamma_\mu \frac{i(\gamma, P_- - k) - m}{D_-} (e\gamma) + (e\gamma) \frac{i(\gamma, k - P_+) - m}{D_+} \gamma_\mu \right] u(-P_+).$$

We shall calculate A_μ separately for each case according to whether the spin of the incident particle flips or not and also whether the polarizations of the virtual photons are transverse or longitudinal. Though this round about procedure seems to be more complicated than the spur-calculation, it makes it much easier to pick out the main terms of $\sum |K|^2$, and we can easily find how much the effects of flipping of the spin of the incident particle and the polarizations of the virtual photon contribute to the cross section.

Now, taking into account (6) and (7) we obtain the expressions for the non-spin flip case :

$$\begin{aligned} A_1(\uparrow\uparrow) &= i\bar{u}_\uparrow(P_2)\gamma_1 u_\uparrow(P_1) = \frac{1}{2} \left(\frac{p^+}{E_1} + \frac{p^-}{E_2} \right), \\ A_2(\uparrow\uparrow) &= i\bar{u}_\uparrow(P_2)\gamma_2 u_\uparrow(P_1) = \frac{i}{2} \left(-\frac{p^+}{E_1} + \frac{p^-}{E_2} \right), \\ A_3(\uparrow\uparrow) &= i\bar{u}_\uparrow(P_2)\gamma_3 u_\uparrow(P_1) = 1 - \frac{\mu^2}{8} \left(\frac{E_1 + E_2}{E_1 E_2} \right)^2 - \frac{p_1^2}{4} \left(\frac{1}{E_1^2} + \frac{1}{E_2^2} \right), \\ A_4(\uparrow\uparrow) &= i\bar{u}_\uparrow(P_2)\gamma_4 u_\uparrow(P_1) = i \left\{ 1 - \frac{\mu^2}{8} \left(\frac{E_1 - E_2}{E_1 E_2} \right)^2 - \frac{p_1^2}{4} \left(\frac{E_1 - E_2}{E_1 E_2} \right)^2 \right\} \\ &= iA_3(\uparrow\uparrow) + \frac{i}{2} \frac{\mu^2 + p_1^2}{E_1 E_2}, \end{aligned} \quad (9)$$

and for the spin flip case :

$$\begin{aligned} A_1(\downarrow\uparrow) &= i\bar{u}_\downarrow(P_2)\gamma_1 u_\uparrow(P_1) = -(\mu/2) (E_1 - E_2)/E_1 E_2, \\ A_2(\downarrow\uparrow) &= i\bar{u}_\downarrow(P_2)\gamma_2 u_\uparrow(P_1) = i(\mu/2) (E_1 - E_2)/E_1 E_2, \\ A_3(\downarrow\uparrow) &= i\bar{u}_\downarrow(P_2)\gamma_3 u_\uparrow(P_1) = -(p^-/2) (E_1 - E_2)/E_1 E_2, \\ A_4(\downarrow\uparrow) &= i\bar{u}_\downarrow(P_2)\gamma_4 u_\uparrow(P_1) = -i(p^-/2) (E_1 - E_2)/E_1 E_2, \end{aligned} \quad (10)$$

where

$$p_{\pm}^{\pm} = p_{\pm}^{\pm} = p_{\pm}^{\pm}, \quad p_{\pm}^{\pm} = p_{1x} \pm ip_{1y}, \quad p_{\pm}^{\pm} = p_{2x} \pm ip_{2y},$$

$$p_{\pm}^0 = p_{\pm}^0 = p_{\pm}^0, \quad p_{\pm}^0 = p_{1x}^0 + p_{1y}^0, \quad p_{\pm}^0 = p_{2x}^0 + p_{2y}^0.$$

For $A_{\mu}(\downarrow\downarrow)$ and $A_{\mu}(\uparrow\downarrow)$ we can obtain similar expressions to (9) and (10).

Inserting (9) and (10) into (8), using the approximation (7) and averaging over φ_2 , we obtain

$$\begin{aligned} \sum |K|^2 = & (4\pi)^4 (Ze^2)^2 \left(\frac{e^2}{m} \right)^2 \frac{1}{q^4 k^4} \frac{\epsilon_+^2 \epsilon_-^2}{D_+^2 D_-^2} \\ & \times \left\{ \left[\left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} \frac{p_{\pm}^2}{\epsilon_+ \epsilon_-} + \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} \frac{\epsilon_-}{\epsilon_-} D_+^2 + \left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} \frac{p_{\pm}^2}{\epsilon_+ \epsilon_-} + \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} \frac{\epsilon_+}{\epsilon_+} D_-^2 \right. \right. \\ & \quad \left. \left. + 2 \left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} (\mathbf{p}_+ \mathbf{p}_-)_{\perp} - \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} \left(\frac{1}{E_1} + \frac{1}{E_2} \right) p_{\pm}^2 \right. \right. \\ & \quad \left. \left. + \left[2 \left\{ \left(\frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \right)^2 - 2 \frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} + \left(\frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} \right)^2 \right\} \frac{\epsilon_-}{\epsilon_-} D_+^2 \right. \right. \right. \\ & \quad \left. \left. + 2 \left\{ \left(\frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \right)^2 - 2 \frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} + \left(\frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} \right)^2 \right\} \frac{\epsilon_+}{\epsilon_+} D_-^2 \right. \right. \\ & \quad \left. \left. - 4 \left\{ \left(\frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \right)^2 - \frac{\mu^2 + p_{\pm}^2}{E_1 E_2} \left(\frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} + \frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} \right) + \frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} \frac{m^2 + p_{\pm}^2}{\epsilon_+ \epsilon_-} \right\} D_+ D_- \right] \right. \\ & \quad \left. + \left[\left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} \frac{p_{\pm}^2}{\epsilon_+ \epsilon_-} + \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} \frac{\epsilon_-}{\epsilon_-} D_+^2 + \left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} \frac{p_{\pm}^2}{\epsilon_+ \epsilon_-} + \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} \frac{\epsilon_+}{\epsilon_+} D_-^2 \right. \right. \\ & \quad \left. \left. + 2 \left\{ \frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon_+ \epsilon_-} (\mathbf{p}_+ \mathbf{p}_-)_{\perp} - \frac{\epsilon^2}{\epsilon_+ \epsilon_-} \frac{m^2}{\epsilon_+ \epsilon_-} \right\} D_+ D_- \right] \left(\frac{\epsilon^2 \mu}{E_1 E_2} \right)^2 \right\}, \quad (11) \end{aligned}$$

where

$$\mathbf{p}_{\pm\perp} = (p_{\pm x}, p_{\pm y}, 0, 0),$$

$$(\mathbf{p}_+ \mathbf{p}_-)_{\perp} = (\mathbf{p}_{+\perp} \mathbf{p}_{-\perp}).$$

The general forms of the terms involved in the outer curly brackets of $\sum |K|^2$ are $\text{const.} \times 0((\mu/E)^{2k}(m/\epsilon)^{2l})$, here k or l is zero or positive integer. In (11) we pick out only the terms satisfying $k+l=2$. By this procedure we can derive an expression which always involves the corresponding main terms to all values of the transferred energy. The three square brackets in (11) correspond to the contributions from the following three processes, respectively; (i) the spin of the incident particle does not flip and the virtual photons are transverse; (ii) the spin of the incident particle does not flip but the virtual photons are longitudinal and scalar ones; (iii) the spin of the incident particle flips and the virtual photons are transverse. The contribution from the process that the spin flips and the virtual photons are longitudinal and scalar ones is neglected on account of the smallness of its order of magnitude. The interference terms between the transverse and the

longitudinal photons vanish by the average over φ_2 . Using the approximation (7), we can write down \mathbf{q}^0 , k^2 and D_{\pm} as follows :

$$\mathbf{q}^0 = (\mathbf{p}_+ + \mathbf{p}_-)_{\perp} + \frac{1}{4} \left\{ \frac{1}{\epsilon_+} (m^2 + p_{+\perp}^2) + \frac{1}{\epsilon_-} (m^2 + p_{-\perp}^2) + \frac{\epsilon_-}{E_1 E_2} (\mu^2 + p_{\perp}^2) \right\}^2,$$

$$k^2 = (\epsilon^2/E_1 E_2) (\mu^2 + p_{\perp}^2), \quad (12)$$

$$D_{\pm} = \frac{\epsilon}{\epsilon_{\pm}} \left(M^2 + P_{\pm\perp}^2 + \frac{\epsilon_+ \epsilon_-}{E_1 E_2} p_{\perp}^2 \right), \quad M^2 \equiv m^2 + (\epsilon_+ \epsilon_-/E_1 E_2) \mu^2$$

(c) Integrations over angles

In order to perform angular integrations, we must transform the density of the final state ρ_F into the S-system (see Appendix). Its form is

$$\rho_F = \frac{m^2}{\epsilon_+ \epsilon_-} \frac{1}{(2\pi)^9} \left(\frac{\epsilon}{E_1} \right)^2 d\mathbf{p}_{+\perp} d\mathbf{p}_{-\perp} d\mathbf{p}_{\perp} d\epsilon_+ d\epsilon_- dE_2, \quad (13)$$

or if we introduce a new set of variables instead of $\mathbf{p}_{\pm\perp}$,

$$\boldsymbol{\zeta} = (\mathbf{p}_+ + \mathbf{p}_-)_{\perp}, \quad \boldsymbol{\eta} = (\mathbf{p}_+ - \mathbf{p}_-)_{\perp}, \quad (14)$$

the final form of ρ_F is given by

$$\rho_F = \frac{m^2}{\epsilon_+ \epsilon_-} \frac{1}{(2\pi)^9} \left(\frac{\epsilon}{E_1} \right)^2 \frac{1}{2^6} d(\zeta^2) d\varphi_{\zeta} d(\eta^2) d\varphi_{\eta} d(p_{\perp}^2) d\varphi_2 d\epsilon_+ d\epsilon_- dE_2, \quad (15)$$

where φ_{ζ} and φ_{η} are azimuthal angles of $\boldsymbol{\zeta}$ and $\boldsymbol{\eta}$, respectively. By the transformation (14), (12) becomes

$$\mathbf{q}^0 = \boldsymbol{\zeta}^2 + \left(\frac{\epsilon}{2\epsilon_+ \epsilon_-} \right)^2 \left[M^2 + \frac{1}{4} (\zeta^2 + \eta^2) + \frac{1}{2} \frac{\epsilon_- - \epsilon_+}{\epsilon} (\boldsymbol{\zeta} \boldsymbol{\eta}) + \frac{\epsilon_+ \epsilon_-}{E_1 E_2} p_{\perp}^2 \right]^2,$$

$$D_{\pm} = \frac{\epsilon}{\epsilon_{\pm}} \left[M^2 + \frac{1}{4} \{ \zeta^2 + \eta^2 \pm 2 (\boldsymbol{\zeta} \boldsymbol{\eta}) \} + \frac{\epsilon_+ \epsilon_-}{E_1 E_2} p_{\perp}^2 \right]. \quad (16)$$

From (16) we can see that the main contribution to the integration over $\boldsymbol{\zeta}$ and $\boldsymbol{\eta}$ comes from the domain where the order of magnitude of η is $0(M)$ and the order of magnitude of ζ is $0(M^2/\epsilon)$ (not $0(M)$). Therefore in the common denominator $1/q^4 D_+^2 D_-^2$ we neglect $\boldsymbol{\zeta}$ in comparison with $\boldsymbol{\eta}$ and p_{\perp} , i.e.,

$$\mathbf{q}^0 = \boldsymbol{\zeta}^2 + \left(\frac{\epsilon}{2\epsilon_+ \epsilon_-} \right)^2 \left[M^2 + \frac{\eta^2}{4} + \frac{\epsilon_+ \epsilon_-}{E_1 E_2} p_{\perp}^2 \right]^2,$$

$$D_{\pm} = \frac{\epsilon}{\epsilon_{\pm}} \left[M^2 + \frac{\eta^2}{4} + \frac{\epsilon_+ \epsilon_-}{E_1 E_2} p_{\perp}^2 \right]. \quad (17)$$

Here the azimuthal angles φ_{ζ} and φ_{η} disappear. The neglect of φ_{ζ} and φ_{η} may be justified by introducing the S-system even when the transferred energy is large. But in the S'-system (see Appendix) used in B we can not neglect φ_{η} , because the expressions corresponding to (16) contain a term like $(\epsilon_+ \epsilon_-/E_1 E_2) (\mathbf{p} \boldsymbol{\eta})_{\perp}$ which can not be neglected

in the case of the large transferred energy. For this reason the calculation procedure using the S' -system will be very complicated for large transferred energy. After averaging over φ_ζ , we pick out only the terms of the lowest order of ζ which are proportional to ζ^2 and obtain the next transformed expression of (11),

$$\begin{aligned} \sum |K|^2 = & (4\pi)^{-1} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \frac{\zeta^2}{q^4} \\ & \times \left[\left\{ \frac{m^2 \epsilon_+ \epsilon_-}{D^4 k^4} \eta^2 - \frac{1}{D^4 k^2} \frac{\epsilon_+ \epsilon_-}{2\epsilon^2} (\epsilon_+^2 + \epsilon_-^2) \eta^2 + \frac{1}{k^4 D^2} (\epsilon_+^2 + \epsilon_-^2) \right\} \left(\frac{1}{E_1^2} + \frac{1}{E_2^2} \right) p_\perp^2 \right. \\ & + \frac{4}{D^4} \left(\frac{\epsilon_+ \epsilon_-}{\epsilon^2} \right) \eta^2 \\ & \left. + \left(\frac{\epsilon \mu}{E_1 E_2} \right)^2 \left\{ \frac{m^2 \epsilon_+ \epsilon_-}{D^4 k^4} \eta^2 - \frac{1}{D^4 k^2} \frac{\epsilon_+ \epsilon_-}{2\epsilon^2} (\epsilon_+^2 + \epsilon_-^2) \eta^2 + \frac{1}{k^4 D^2} (\epsilon_+^2 + \epsilon_-^2) \right\} \right], \quad (18) \end{aligned}$$

where

$$D \equiv M^2 + \eta^2/4 + (\epsilon_+ \epsilon_- / E_1 E_2) p_\perp^2. \quad (19)$$

By the above transformation, integrations over the angles have been replaced by ζ^2 , η^2 and p_\perp^2 . Since (18) has been averaged over the variables φ_η , φ_η and φ_ζ , integrations over these variables will give only $(2\pi)^3$. We take the integration domain of ζ^2 from zero to $(\alpha M)^2$, where α is a number of order unity, and for η^2 and p_\perp^2 from zero to infinity as a good approximation.

When the incident particle collides with a neutral atom, it is necessary to consider the screening effect. This effect may be easily taken into account if we replace ζ^2/q^4 in (18) by

$$\zeta^2/q^4 \cdot [1 - F(q^2)]^2, \quad (20)$$

where $F(q^2)$ is the atomic form factor. The minimum value of q^2 is given from (17) by

$$q_{\min}^2 = (\epsilon/2\epsilon_+ \epsilon_-)^2 M^4, \quad (21)$$

hence the screening is effective when

$$(\epsilon/2\epsilon_+ \epsilon_-) M^2 \ll Z^{1/3} m/137. \quad (22)$$

In this case, integration over ζ^2 can be done in the same manner as in B. As to pair creation by γ -rays, more precise calculations of the screening effect were made by Bethe¹¹, but since the treatment of B is in principle the same as Bethe's calculation, we will follow B.

(d) Differential cross section

Inserting (17), (18), (19), and (15) into (5), and after performing the integrations over ζ^2 , η^2 , p_\perp^2 , φ_η , φ_ζ and $\varphi_{r\eta}$, we get finally, as the differential cross section for the creation of a pair, the electron of which has an energy between ϵ_- and $\epsilon_- + d\epsilon_-$, and the positron of which has an energy between ϵ_+ and $\epsilon_+ + d\epsilon_+$,

$$\begin{aligned}
\sigma = & 2/\pi \cdot (Ze^2)^2 (e^2/m)^2 d\epsilon_+ d\epsilon_- L \\
& \times \left\{ \left[\frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon^4} \left\{ \left(1 + \frac{4}{3}x \right) \log \left(1 + \frac{1}{x} \right) - \frac{4}{3} \right\} + \frac{2}{3} \frac{\epsilon_+ \epsilon_-}{\epsilon^4} \left\{ (1+2x) \log \left(1 + \frac{1}{x} \right) - 2 \right\} \right] \right. \\
& \times \frac{E_1^2 + E_2^2}{E_1^2} \\
& + \frac{8}{3} \frac{\epsilon_+ \epsilon_-}{\epsilon^4} \frac{1}{1+x} \frac{E_2}{E_1} \\
& + \left. \left[\frac{\epsilon_+^2 + \epsilon_-^2}{\epsilon^4} \left\{ \frac{1}{3} \frac{1}{1+x} + \frac{1}{x} - \frac{4}{3} \log \left(1 + \frac{1}{x} \right) \right\} \right. \right. \\
& \left. \left. + \frac{2}{3} \frac{\epsilon_+ \epsilon_-}{\epsilon^4} \left\{ \frac{1}{1+x} + \frac{1}{x} - 2 \log \left(1 + \frac{1}{x} \right) \right\} \right] x \frac{\epsilon_-^2}{E_1^2} \right\}, \quad (23)
\end{aligned}$$

where

$$x = (\mu^2/E_1 E_2) \cdot (\epsilon_+ \epsilon_-/m^2), \quad (24)$$

and

$$L = \begin{cases} \log(2\alpha\epsilon_+\epsilon_-/\epsilon M) - 1 & \text{for non-screening,} \\ \log\{\alpha \cdot 137 Z^{-1/3} \cdot M/m\} & \text{for complete screening.} \end{cases} \quad (25)$$

In the expression (23), the terms proportional to $(E_1^2 + E_2^2)/E_1^2$ and E_2/E_1 correspond respectively to the process in which the spin of the incident particle does not flip and the direction of polarization of the virtual photon is transverse, and to the process in which the spin of the incident particle does not flip but the direction of polarization of the virtual photon is longitudinal. The term proportional to ϵ_-^2/E_1^2 corresponds to the process that the spin flips and the virtual photon polarizes transversely.

In order to get approximate expressions to (23) for each of both cases that the transferred energy is large or small, it is convenient to introduce a set of quantities instead of ϵ_{\pm} , i.e.,

$$\begin{aligned}
u &= (\epsilon_+ + \epsilon_-)/E_1 = \epsilon/E_1, \\
v &= (\epsilon_+ - \epsilon_-)/\epsilon. \quad (26)
\end{aligned}$$

We say the transferred energy is small when

$$u \ll m/\mu, \quad (27)$$

and large when

$$u \gg m/\mu. \quad (28)$$

Hereafter we shall call the former the domain I and the latter II. If we express x by u and v , we obtain from (24) and (26)

$$x = \frac{1}{4} (u\mu/m)^2 \cdot (1-v^2)/(1-u). \quad (29)$$

Since in the domain I always $x \ll 1$, we obtain as the approximate formula of the differential cross section in this domain

$$\sigma_1 = \frac{4}{\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \frac{d\epsilon_+ d\epsilon_-}{\epsilon^4} \left[\log \frac{2\alpha\epsilon_+ \epsilon_-}{\epsilon m} \right] (\epsilon_+^2 + \epsilon_-^2 + \frac{2}{3}\epsilon_+ \epsilon_-) \log \frac{m^2 E_1^2}{\mu^2 \epsilon_+ \epsilon_-} \quad (30)$$

$$= \frac{4}{3\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \frac{dudv}{u} \left[\log \frac{\alpha E_1}{2m} u(1-v^2) \right] \left(1 + \frac{v^2}{2} \right) \log \left\{ \left(\frac{m}{u\mu} \right)^2 \frac{4}{1-v^2} \right\}. \quad (31)$$

If we replace the factor $\log(m^2 E_1^2 / \mu^2 \epsilon_+ \epsilon_-)$ in (30) by $\log(k'^2 m^2 E_1^2 / \mu^2 \epsilon^2)$ where k' is a number of order unity, (30) is the same as that of B, but the indefinite k' has disappeared in our calculations. As was expected, (30) corresponds only to the process that the spin does not flip and the direction of the polarization of the virtual photon is transverse.

In the domain II, it will be necessary to note that there are two regions according to whether or not the energy transferred to the electron is almost equal to that transferred to the positron because from (26) and (29)

$$x \gg 1 \quad \text{for} \quad v^2 \ll 1, \quad (32)$$

$$x \ll 1 \quad \text{for} \quad 1 - v^2 \ll 1. \quad (33)$$

The approximate formulas in these two domains are written as

$$\begin{aligned} \sigma_{IIa} &= \frac{2}{3\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \left(\frac{m}{\mu} \right)^2 \frac{d\epsilon_+ d\epsilon_-}{\epsilon^4} \frac{1}{x} L \\ &\quad \times \left[(\epsilon_+^2 + \epsilon_-^2) \left(1 + \frac{E_2^2}{E_1^2} \right) + 8\epsilon_+ \epsilon_- \frac{E_2}{E_1} + 2(\epsilon_+^2 + \epsilon_-^2 + \epsilon_+ \epsilon_-) \left(\frac{\epsilon}{E_1} \right)^2 \right] \\ &= \frac{4}{3\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \left(\frac{m}{\mu} \right)^2 \frac{dudv}{u^3} \frac{1-u}{1-v} \log \left[\frac{\alpha E_1}{\mu} (1-v^2)^{1/2} (1-u)^{1/2} \right] \\ &\quad \times \left[(3-v^2) \left(1 + \frac{u^2}{4} \right) - 3u + \frac{1}{2} u^2 (1+v^2) \right], \end{aligned} \quad (34a)$$

and

$$\begin{aligned} \sigma_{IIb} &= \frac{4}{\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \left(\frac{m}{\mu} \right)^2 \frac{d\epsilon_+ d\epsilon_-}{\epsilon^4} \left[\log \frac{2\alpha\epsilon_+ \epsilon_-}{\epsilon m} - 1 \right] \\ &\quad \times \left[\left(1 - \frac{\epsilon}{E_1} + \frac{\epsilon^2}{E_1^2} \right) (\epsilon_+^2 + \epsilon_-^2 + \frac{2}{3}\epsilon_+ \epsilon_-) \log \frac{m^2 E_1 E_2}{\mu^2 \epsilon_+ \epsilon_-} \right. \\ &\quad \left. - \frac{4}{3} (\epsilon_+^2 + \epsilon_-^2) \left(1 + \frac{\epsilon}{E_1} \right) - \frac{4}{3} \left(\frac{\epsilon}{E_1} \right)^2 (\epsilon_+^2 + \epsilon_-^2 + \epsilon_+ \epsilon_-) \right] \\ &= \frac{4}{3\pi} (Ze^0)^2 \left(\frac{e^2}{m} \right)^2 \left(\frac{m}{\mu} \right)^2 \frac{dudv}{u} \left[\log \frac{\alpha E_1 u (1-v^2)}{2m} - 1 \right] \\ &\quad \times \left[\left(1 - u + \frac{u^2}{2} \right) \left(1 + \frac{v^2}{2} \right) \log \left\{ \left(\frac{2m}{u\mu} \right)^2 \frac{1-u}{1-v^2} \right\} - (1+v^2)(1-u) - \frac{1}{4} (3+v^2) u^2 \right]. \end{aligned} \quad (34b)$$

We can also obtain in the similar manner the completely screened differential cross section.

§ 3. Total cross section

(a) Non-screened total cross section

To get the contribution to the total cross section from the domain I, we take the integral domains of u and v as follows,

$$0 \leq u \leq m/\mu, \quad -1 + m/\epsilon \leq v \leq 1 - m/\epsilon. \quad (35)$$

Though the expression (31) does not hold at the two limits of integration, (35), u and v occur only in logarithmic functions after the integration of (31) so that the indefiniteness of the domains of integration of u and v does not appreciably affect the final results. The result of the integration of (31) is

$$Q_I^n = 28/(27\pi) \cdot (Ze^2)^2 (e^2/m)^2 (\log \alpha E_1/\mu)^3. \quad (36)$$

We have neglected the terms of lower power of $\log(E_1/\mu)$ in (36), because the errors which come from the above mentioned indefiniteness of the integration domains are of the order of $(\log E_1/\mu)^2$.* Similarly we obtain in the domain IIa and IIb, respectively,

$$Q_{IIa}^n = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 (\log 2 + 2) \log \frac{\alpha E_1}{\mu}, \quad (37a)$$

$$Q_{IIb}^n = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left(\frac{3}{2} \log 2 - \frac{1}{2}\right) \log \frac{\alpha E_1}{\mu}, \quad (37b)$$

where β is the number which occurs from the limit of the integration over v , i.e., $\beta(m/\mu) \leq u \leq 1$. The sum of the above two values (37a) and (37b) gives the contribution from the domain II to the non-screened total cross section, i.e.,

$$Q_{II}^n = Q_{IIa}^n + Q_{IIb}^n = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left(\frac{5}{2} \log 2 + \frac{3}{2}\right) \log \frac{\alpha E_1}{\mu}. \quad (38)$$

It must be noted that Q_{II}^n depends on the inverse square of β^2 , hence the indefiniteness of β seriously affects the final result. This point will be discussed again in the last section.

(b) Completely screened total cross section

In this case each domain consists of two parts, one of which satisfies (22) and the other does not. The contribution to the cross section from each domain is given by the sum of the contributions from the two parts. The results are as follows:

* Including the terms of lower power of $\log(E_1/\mu)$, the full expression of (36) is written as

$$Q_I^n = \frac{4}{3\pi} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left[\frac{7}{9} \left(\log \frac{E_1}{\mu}\right)^3 + C_1 \left(\log \frac{E_1}{\mu}\right)^2 + C_2 \log \frac{E_1}{\mu} \right],$$

where $C_1 = (4/3) \log \alpha - 0.85$, $C_2 = 20.81 - (14/3) (\log \beta_1)^2 - (62/9) \log \alpha + (28/3) \log(3/4) \log \beta_2$, and β_1 and β_2 are the numbers ($\gtrsim 1$) which occur from the limits of the integration.

$$Q_i^s = \frac{28}{27\pi} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \log(137 Z^{-1/3}) \left[3 \log \frac{\alpha E_1}{\mu} \log \frac{E_1}{137 Z^{-1/3} \mu} + (\log 137 Z^{-1/3})^2 \right], \quad (39)$$

$$Q_{ii}^s = Q_{iia}^s + Q_{iib}^s = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left(\frac{5}{2} \log 2 + \frac{3}{2} \right) \log(\alpha \cdot 137 Z^{-1/3}), \quad (40)$$

where

$$Q_{iia}^s = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 (\log 2 + 2) \log\left(\frac{\alpha}{2} 137 Z^{-1/3}\right), \quad (41a)$$

$$Q_{iib}^s = \frac{4}{3\pi} \frac{1}{\beta^2} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left(\frac{3}{2} \log 2 - \frac{1}{2}\right) \log(\alpha \cdot 137 Z^{-1/3}). \quad (41b)$$

§ 4. The case in which the incident particle is an electron

Now the cross section consists of three parts proportional respectively to $|M_{ii}|^2$, $|M_{ii'}|^2$ and $|M_{ii}M_{ii'}^* + M_{ii'}^*M_{ii}|$. Until now, we have neglected the latter two parts. This is permitted if the rest mass, μ , of the incident particle is much larger than that of an electron, because the part proportional to $|M_{ii'}|^2$ contains the factor $(e^2/\mu)^2$ instead of $(e^2/m)^2$ as in (23) and the interference part proportional to $|M_{ii}M_{ii'}^* + M_{ii'}^*M_{ii}|$ contains the factor $(e^2/m)(e^2/\mu)$.

However, when the incident particle is an electron, the above discussion is no longer valid and moreover the exchange effect must be considered. The part proportional to $|M_{ii'}|^2$ can be evaluated in a manner similar to that used for the part proportional to $|M_{ii}|^2$ and is given by

$$\begin{aligned} \sigma' = & \frac{2}{\pi} (Ze^2)^2 \left(\frac{e^2}{\mu}\right)^2 d\epsilon_+ d\epsilon_- L' \frac{\epsilon_+^2 + \epsilon_-^2}{E_1^2} \\ & \times \left[\frac{E_1^2 + E_2^2}{\epsilon_+ \epsilon_-} \left\{ \left(1 + \frac{4}{3} x'\right) \log\left(1 + \frac{1}{x'}\right) - \frac{4}{3} \right\} - \frac{2}{3} \frac{E_1 E_2}{\epsilon_+ \epsilon_-} \left\{ (1 + 2x') \log\left(1 + \frac{1}{x'}\right) - 2 \right\} \right], \end{aligned} \quad (42)$$

where

$$x' = 1/x,$$

$$L' = \begin{cases} \log(2\alpha E_1 E_2 / \epsilon \mu (1+x')^{1/2}) - 1 & \text{for non-screening.} \\ \log\{\alpha \cdot 137 Z^{-1/3} (1+x')^{1/2}\} & \text{for complete screening.} \end{cases}$$

As is expected, (42) is a type of Bremsstrahlung.

Since $\mu = m$, the domain II disappears and the only remaining domain is I. The contribution to the non-screened total cross section is evaluated approximately.

$$Q_{ii'} \sim (Ze^2)^2 (e^2/m)^2 \log(E_1/m). \quad (43)$$

The direct evaluation of the interference parts is rather complicated, but its magnitude can easily be estimated as follows:

$$\begin{aligned}
 d\sigma_{DD'} &= \frac{1}{v} |M_D M_{D'}^* + M_{D'}^* M_D| \rho_F \\
 &\leq \frac{2}{v} |M_D| \cdot |M_{D'}| \rho_F = 2 \sqrt{\frac{1}{v} |M_D|^2 \rho_F \cdot \frac{1}{v} |M_{D'}|^2 \rho_F} \\
 &= 2 \sqrt{d\sigma_D d\sigma_{D'}},
 \end{aligned} \tag{44}$$

where v is the velocity of the incident electron. After integration of (44)

$$Q_{DD'} \leq 2 \int \sqrt{d\sigma_D d\sigma_{D'}} \leq 2 \sqrt{\int d\sigma_D \cdot \int d\sigma_{D'}} = 2 \sqrt{Q_D Q_{D'}}. \tag{45}$$

The right hand side is then proportional to $(Ze^2)^2 (e^2/m)^2 (\log E_1/m)^2$ which is of the same order as that of the terms neglected in the evaluation of (36).

Next we shall estimate the exchange effect. We represent by M_D^e the matrix element in which P_2 and P_- are interchanged in M_D and put

$$M_D = \bar{M}_D/k^2, \quad M_D^e = \bar{M}_D^e/k'^2, \quad k' = P_1 - P_-. \tag{46}$$

Then the cross section is given by

$$\sigma \sim \frac{1}{2} |M_D - M_D^e|^2 = \frac{1}{2} \frac{1}{k^4} |\bar{M}_D|^2 + \frac{1}{2} \frac{1}{k'^4} |\bar{M}_D^e|^2 - \frac{1}{k^2 k'^2} \text{Re}(\bar{M}_D^* \bar{M}_D^e). \tag{47}$$

The contribution of the second term to the total cross section is the same as that of the first, and the third term can be estimated by comparing $1/k^2 k'^2$ with $1/k^4$. Thus, putting

$$I_D = \int \frac{1}{k^4} d\Omega_2 d\Omega_-, \quad I_{DD'} = \int \frac{1}{k^2 k'^2} d\Omega_2 d\Omega_-, \tag{48}$$

we get

$$I_{DD'}/I_D < (m/E_1)^2 (\log 2E_1/m)^2, \tag{49}$$

therefore the exchange effect can be neglected as long as $E_1 \gg m$.

§ 5. Summary and discussions

First we shall confine our discussions to the case in which the incident particle is an electron. Since the domain II disappears for $\mu=m$, the total cross section is given by (39) or (36) according to whether the screening is effective or not. In (39) and (36) the terms of lower power of $\log(E_1/m)$ are neglected. This is justified if the energy of the primary electron is larger than about 10 Bev. For example the full expression for (36) is given by

$$Q_1^e = \frac{4}{3\pi} (Ze^2)^2 \left(\frac{e^2}{m}\right)^2 \left[\frac{7}{9} \left(\log \frac{E_1}{m}\right)^3 - C_1 \left(\log \frac{E_1}{m}\right)^2 + C_2 \log \frac{E_1}{m} \right],$$

where $C_1 < 0.85$ and $C_2 < 20.81$.* Thus the second and the third terms can be neglected

* See the footnote on page 491.

if the incident electron has an energy larger than the above mentioned value. The error due to this neglect is at most 20%. For such high energies the contribution from the diagrams D' may also be neglected as estimated in § 4. If the energy of the incident electron is smaller than 10 Bev, we can not disregard the contribution of the diagrams D' and the terms of low power.

As is well known, the cross section derived with the Born approximation is inclined to become larger than the actual value. This error is estimated by Bethe et al.⁶⁾ in the case of the pair creation by γ -rays using the distorted electron wave functions. They showed that the value derived with the Born approximation is over-estimated by a factor of 20%. Since (36) is essentially the same as that derived by the Williams-Weizsäcker method⁷⁾, the same situation will hold in our case. Another effect of suppression has been pointed out by Landau and Pomeranchuk^{8)*} in the cases of Bremsstrahlung and pair creation by γ -rays. This effect arises from the fact that the incident particle collides with a "medium", not with one isolated atom. This suppression becomes important at energies higher than the critical value $\sim 10^{15}$ eV for lead. Though the present processes are not the ones discussed by them, it is certain that this effect also suppresses our values of the cross sections, because this suppression effect is essentially caused by the interference of many waves with different phases. It is noted that for the above two reasons our results (36) and (39) may be considered slightly larger than the actual one.

The experimental values measured by Koshiha and Kaplon may be compared with our results since the energies of primary electrons are larger than 10 Bev. Our cross sections give about one third of their values. This discrepancy may not be considered conclusive, because the experimental errors due to the measurement of the energies of the primary particles are suspected to be large in such a high energy region. Block et al.⁹⁾ showed that their experimental results are consistent with B modified so as to include the terms of lower power of $\log(E_1/m)$. However, since in their experiment the primary energies lie between 0.1 and 10 Bev having the average value 400 Mev, one has to take into account the contributions not only from the terms of lower power but also from the diagrams D' . Otherwise, the comparison with theory is of little meaning. It is unfortunate that a decisive conclusion can not be drawn from the above mentioned experiments. Since we have derived the cross sections for this process using the current theory and taking into account the various effects for high energy, it is desired that the trident process, in which the primary electron has an energy higher than 10 Bev, will be studied more extensively.

In the case that the mass of the incident particle is heavy compared with that of an electron, the differential cross section is given by (23), which is valid as long as the participant particles have relativistic energies and is more correct than those considered heretofore. As to small transferred energies the approximation formula (30) to (23) is essentially the same as B. However for large transferred energies there are two differences

* The authors thank sincerely to Prof. Z. Koba who kindly brought Landau and Pomeranchuk's work to their notice.

between our calculation and that of B, one of which is the crudeness in the estimation of the domain IIa in B*, the other being the fact that the domain IIb is not discussed in B.

Now the total cross section is the sum of the contributions from the domain I and II, i.e., $Q_I^n + Q_{II}^n$ for non-screening and $Q_I^s + Q_{II}^s$ for complete screening. Q_{II}^n (or Q_{II}^s) depends on the indefinite number β^{**} ; the maximum estimate of Q_{II}^n (or Q_{II}^s) is given by putting $\beta=1$. But in any case we may neglect Q_{II}^n (or Q_{II}^s) compared with Q_I^n (or Q_I^s). For the energy loss of the incident particle we cannot neglect the contribution from domain II, which will be of the same order as that from domain I. Thus we see that the final results seriously depend on the value of β . Therefore the value of the energy loss using B may have an error of a factor about two.***

Finally, we wish to make one more remark. It was shown that the main contribution to the process comes from the transition in which the spin of the incident particle does not flip and the virtual photon polarizes transversely. This fact seems to show that the Williams-Weizsäcker method should be a good approximation. We shall discuss this point in a forthcoming paper.

The authors wish to express their cordial thanks to Prof. T. Inoue for his encouragement, Dr. H. Hasegawa for his valuable discussions and Mr. Edo for his assistance in carrying out the preliminary calculations.

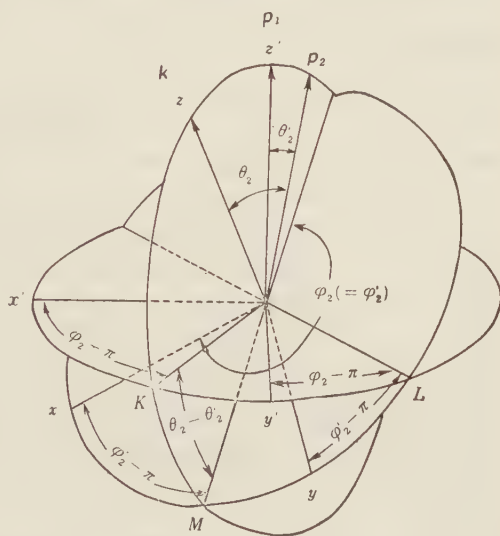
Appendix

We shall call a coordinate system, in which the z' -axis is parallel to \mathbf{p}_1 , the S' (x', y', z')-system, and a coordinate system having its z -axis parallel to $\mathbf{k} = \mathbf{p}_1 - \mathbf{p}_2$ the S (x, y, z)-system. We take the x - and y -axes of the S -system as shown in fig. A1.

The density of the final state in the S' -system is given by

$$\rho_F' = d\mathbf{p}_2' / (2\pi)^3 \cdot (m^2 / \epsilon_+ \epsilon_-) \cdot$$

$$\cdot d\mathbf{p}_+ d\mathbf{p}_- / (2\pi)^6,$$



$$\begin{array}{ccc} x' y' z' & \xrightarrow{\varphi_2 - \pi} & KL z' \xrightarrow{\theta_2 - \theta_2'} ML z \\ & \xrightarrow{-(\varphi_2' - \pi)} & xyz \end{array}$$

Fig. A1.

* In order to derive (34a), we used an approximation: $\log\left(1 + \frac{1}{x}\right) = \frac{1}{x} - \frac{1}{2x^2}$, but if we use such a rough approximation as $\log\left(1 + \frac{1}{x}\right) = \frac{1}{x}$, we obtain the same expression as derived in B (see eq. (36) in B).

** In B, β is taken as 3-4. Putting this value into our expressions, our results become the same as B.

*** For large transferred energies there may arise the effect of virtual meson cloud of the target nucleus. Considering the indefiniteness of β , however, our result will not be changed largely by this effect.

where

$$d\mathbf{p}'_2 = dp'_{2x} dp'_{2y} dp'_{2z} = p_2^2 dp_2 \sin \theta'_2 d\theta'_2 d\varphi'_2,$$

$$p_2 = |\mathbf{p}'_2| = |\mathbf{p}_2|.$$

We shall express it by the variables in the S-system. The form of the density with respect to the electron pair is not changed and expressed as $(m^2/\epsilon_+ \epsilon_-) \cdot (d\mathbf{p}_+ d\mathbf{p}_- / (2\pi)^6)$, because the transformation between the two systems is orthogonal. With respect to the incident charged particle we obtain from figs. A1 and A2,

$$\varphi'_2 = \varphi_2,$$

$$p_1 \sin \theta_2 = |\mathbf{k}| \sin \theta'_2,$$

or, using the small angle approximation,

$$\theta'_2 = \epsilon/E_1 \cdot \theta_2.$$

Therefore

$$d\mathbf{p}'_2 = p_2^2 dp_2 \sin \theta'_2 d\theta'_2 d\varphi'_2 \cong (\epsilon/E_1)^2 p_2^2 dp_2 \theta_2 d\theta_2 d\varphi_2 = (\epsilon/E_1)^2 d\mathbf{p}_{2\perp} dp_2,$$

and the density of the final state in the S-system is expressed as

$$\begin{aligned} \rho_F &= \frac{1}{(2\pi)^3} \left(\frac{\epsilon}{E_1} \right)^2 d\mathbf{p}_{\perp} dp_2 \frac{m^2}{\epsilon_+ \epsilon_-} \frac{d\mathbf{p}_+ d\mathbf{p}_-}{(2\pi)^6} \\ &\cong \frac{m^2}{\epsilon_+ \epsilon_-} \frac{1}{(2\pi)^9} \left(\frac{\epsilon}{E_1} \right)^2 d\mathbf{p}_{+\perp} d\mathbf{p}_{-\perp} dp_{+\perp} d\epsilon_+ d\epsilon_- dE_2. \end{aligned}$$

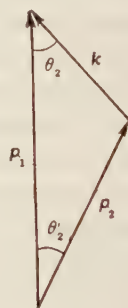


Fig. A2.

It must be noted that B_u does not depend on φ_2 , because B_u depends on \mathbf{p}'_2 only through the momentum $\mathbf{k} (= \mathbf{p}'_1 - \mathbf{p}'_2 = \mathbf{p}_1 - \mathbf{p}_2)$ which is independent of φ_2 . Therefore we may take into consideration A_μ only, when we integrate or average over φ_2 .

References

- 1) H. J. Bhabha, Proc. Roy. Soc. A **152** (1935), 559, quoted as B in our paper.
H. J. Bhabha, Proc. Camb. Phil. Soc. **31** (1935), 394.
- 2) Y. Nishina, S. Tomonaga and M. Kobayasi, Sci. Pap. Inst. Phys. Chem. Research, Japan **27** (1935), 137.
E. J. Williams, Kgl. Dansk. Vid. Selsk. **13** (1935), No. 4.
- 3) S. Hayakawa, Soryushi-ron no Kenkyu, No. 3, (1951), 57 (in Japanese).
- 4) M. Koshiba and M. F. Kaplon, Phys. Rev. **97** (1955), 193; *ibid.*, **100** (1955), 327.
- 5) H. A. Bethe, Proc. Camb. Phil. Soc. **30** (1934), 524.
- 6) H. A. Bethe and L. C. Maximon, Phys. Rev. **93** (1954), 768.
Handel Davies, H. A. Bethe and L. C. Maximon, *ibid.*, **93** (1954), 788.
- 7) E. J. Williams, *loc. cit.*
C. F. v. Weizsäcker, Zeit. f. Phys. **88** (1934), 612.
- 8) L. Landau and I. Pomeranchuk, Doklady Akademii Nauk SSSR **92** (1953), 535, 735.
- 9) M. M. Block, D. T. King and W. W. Wada, Phys. Rev. **96** (1954), 1627.

On the Foundation and the Applicability of Williams-Weizsäcker Method

Toshiyuki MUROTA and Akira UEDA

Department of Physics, Kyoto University, Kyoto

(Received July 9, 1956)

The energy spectrum of the virtual photon field accompanied by a high energy charged particle which has been derived classically by Williams and Weizsäcker independently is derived here by the Feynman-Dyson method, and the equivalence between the Williams-Weizsäcker and the Feynman-Dyson methods is verified in a general way when the energy transferred from an incident charged particle is small. It is also qualitatively shown that the cross section derived by the Williams-Weizsäcker method gives a correct order of magnitude.

§ 1. Introduction

The computation of radiative processes occurring in high energy collisions often turns out to be prohibitively complicated. To avoid this complication Williams¹⁾ and Weizsäcker²⁾ proposed independently an approximation method based on the following idea. Let us consider a fast charged particle moving with velocity almost equal to light velocity. The field of this particle is almost identical with that of a set of photons with various frequencies. The electromagnetic action of this particle on another charged particle, say at rest, is equivalent to that of these virtual photons. This simplified method (W-W method) is often of great help and yields the same results as the Feynman-Dyson method (F-D method) in several cases, for example Bremsstrahlung³⁾, the pair production by a charged particle⁴⁾ and the π production by μ -meson⁵⁾, etc. This agreement is based on the fact that the longitudinal and scalar photon fields accompanied by the incident charged particle is neglected as compared with the transverse one as was shown classically by Williams and also implicitly on the fact that the energy spectrum of the virtual transverse photon derived by him is essentially the same as the one to be derived from the F-D method. In our previous paper⁶⁾ on the electron pair creation by charged particles, we showed that the contribution to the cross section arises mainly from the process that the spin of the incident particle does not flip and the direction of polarization of the virtual photon is transverse, and the energy spectrum of the virtual transverse photon becomes the same as the classical one when the energy transferred from the incident particle to the created pair is very small. This conclusion was derived essentially by separating the contributions to the cross section according to the directions of the polarization of the virtual photon. Therefore this separation procedure suggests the possibility of extending the above conclusion to more general cases. To generalize this conclusion from the quantum electrodynamical point of view is the content of this paper.

In next section we shall review the assumptions and procedures used in the W-W method and pick out the points to be verified for showing the equivalence between the W-W method and the F-D method. Then we shall, in § 3, re-formulate the transition probability by the same procedure as I. In § 4 we shall first show that the energy spectrum of the virtual transverse photon coincides with the classical one and secondly derive the condition that the contributions from longitudinal and scalar photons can be neglected as compared with that from transverse ones. In the final section we shall discuss our assumption introduced in § 3 and make some remarks about our treatment.

§ 2. Review of the W-W method

We shall review briefly the assumptions and procedures which are used in the W-W method. The basic assumptions are as follows: (a) The path of a fast particle remains practically straight during a collision; and (b) the effects of the different Fourier amplitudes in the virtual photon field can be considered as independent. The assumption, (a), requires that the energy transferred by the fast particle, i.e. the momentum and the energy of the virtual photon, is much smaller than that of the fast particle itself. There may be effects which are due to interference between virtual photons of different momenta and different directions of polarization. The former which arises from higher order radiative corrections is neglected in our procedure, and the latter, owing to different polarizations of the transverse photons, will disappear after averaging over the azimuthal angle, φ_2 , of the fast particle in the final state. Under these two assumptions the energy spectrum of the virtual photon can be expressed as*

$$q(\epsilon)d\epsilon = \frac{2}{\pi}(Ze)^2 \frac{d\epsilon}{\epsilon} \log \frac{\alpha m E_1}{\epsilon \mu}, \quad (1)$$

where E_1 is the energy of the fast particle, μ and m are the masses respectively of the fast particle and one of the created particles** in the final state, and α is a constant of order unity. Using this spectrum the cross section of the process in question is written down as

$$Q = \int q(\epsilon) d\epsilon \sigma_T(\epsilon). \quad (2)$$

In this expression $\sigma_T(\epsilon)$ is the cross section caused by a real photon with energy ϵ , which has a form

$$\sigma_T(\epsilon) d\epsilon = \int \frac{1}{2} \sum_{\text{spin}} \frac{1}{2\epsilon} \{ |B_1^\gamma|^2 + |B_2^\gamma|^2 \} \rho'_{\mathbf{p}}, \quad (3)$$

* In Williams' paper $\log(\alpha E_1/\epsilon)$ is used instead of $\log(\alpha m E_1/\epsilon \mu)$. For this point see Bhabha's paper (Proc. Roy. Soc. A, 152 (1935), 559). However this difference can be neglected for $E_1/\epsilon \gg \mu/m$.

** In our sense "particles created in the final state" means not only the particles actually created but also the target particle, except for the case that the action of the target may be replaced by that of an external field. The word "created particle" is used in this sense in what follows.

where \int means the integration over the final density ρ'_F and $1/2\epsilon$ arises from the normalization of the incident photon waves. B_1^τ and B_2^τ correspond respectively to the matrix elements for the cases in which the directions of polarization of the incident photon are parallel to the x - and y -axes if the z -axis of the coordinate system is chosen to be parallel to the direction of propagation of the incident photon. The factor $1/2$ arises from averaging over the directions of polarization.

In the F-D method the matrix element M is the product of two factors, one of which, A_μ , is related with the incident particle, and the other, B_μ , with the process caused by the virtual photon, i.e.,

$$M = A_\mu B_\mu \quad (4)$$

The transition probability may be written as

$$\sum |M|^2 \rho_F = \sum |A_1 B_1 + A_2 B_2 + A_3 B_3 + A_4 B_4|^2 \rho_F, \quad (5)$$

where ρ_F , the density of the final state, is related to ρ'_F in (3) as follows:

$$\rho_F = \rho'_F. \quad (6)$$

Here ρ is the density depending only on the incident particle, i.e., $d\mathbf{p}_2/(2\pi)^3$.

In the expressions (4) and (5) the indices, μ do not necessarily correspond to the directions of polarization of the virtual photon. However if we represent (4) and (5) in the S -coordinate system, whose z -axis is parallel to the direction of propagation of the virtual photon,* the indices, μ , should correspond exactly to the direction of polarization. Namely, (5) may be expressed as

$$\sum |M|^2 \rho_F \rightarrow \sum [|B_1|^2 + |B_2|^2] |A_1|^2 + |A_3 B_3 + A_4 B_4|^2 \rho_F. \quad (7)$$

Thus the W-W method requires the following relation to hold under the two assumptions (a) and (b),

$$\int \frac{1}{2} \sum_{\text{spin}} \frac{1}{2\epsilon} \{ |B_1^\tau|^2 + |B_2^\tau|^2 \} \rho'_F d\epsilon q(\epsilon) \simeq \frac{1}{v} \int_{\text{spin}} \{ |B_1|^2 + |B_2|^2 \} |A_1|^2 \rho'_F, \quad (8)$$

where v , the velocity of the incident particle, is nearly equal to light velocity.**

In the next section we shall derive (7).

§ 3. Preliminary formulation

Now we will use the following assumption (quoted as A): The energies of all particles involved, excluding the target particle in the initial state, are large compared with their rest masses and are of the same order as the transferred energy, i.e., the energy of the virtual photon.

* A coordinate system whose z -axis is parallel to the direction of the incident particle will be called S' -system.

** Strictly speaking, v in the S -system is different from one in the S' -system, but this difference can be neglected.

As is easily seen, the corresponding diagram in the W-W method is of only one type as shown in Fig. 1. For the moment then, let us consider only this diagram. As for the diagrams not considered here, we shall discuss them in the last section. From Fig. 1 the matrix element has the form,

$$\begin{aligned}
 M &= \frac{Ze\sqrt{8\pi\epsilon}}{k^2} A_\mu \frac{1}{\sqrt{2\epsilon}} B_\mu \\
 &= \frac{Ze\sqrt{8\pi\epsilon}}{k^2} \left[\frac{1}{\sqrt{2\epsilon}} \{A_1 B_1 + A_2 B_2\} \right. \\
 &\quad \left. + \frac{1}{2\sqrt{2\epsilon}} \{B_3(A_3 + iA_4) + A_3(B_3 + iB_4)\} \right. \\
 &\quad \left. - \frac{i}{2\sqrt{2\epsilon}} \{B_4(A_3 + iA_4) + A_4(B_3 + iB_4)\} \right], \quad (9)
 \end{aligned}$$

where

$$\begin{aligned}
 A_\mu &= \bar{u}(P_2) \gamma_\mu u(P_1), \\
 \epsilon &= E_1 - E_2. \quad (10)
 \end{aligned}$$

We will represent (9) in the S-system and further introduce the small angle approximation, which is justified by the denominators of the propagators in the matrix elements. For example,

$$p_1 \sin \theta_1 \cong p_1 \theta_1, \quad p_1 \cos \theta_1 - E_1 \cong -\frac{1}{2E_1} (\mu^2 + (p_1 \theta_1)^2). \quad (11)$$

It must be noted that the magnitude of the transverse component of each momentum is of the same order as each rest mass. After proceeding as in I,⁶ that is averaging and summing over the directions of the spin and averaging over φ_2 (see Appendix in I), we obtain

$$\begin{aligned}
 \sum_{\text{spin}, \varphi_2} |M|^2 &= 8\pi\epsilon (Ze)^2 / k^4 \\
 &\cdot \left[\frac{1}{2} \{|B_1|^2 + |B_2|^2\} \frac{1}{2\epsilon} \left(\frac{1}{E_1^2} + \frac{1}{E_2^2} \right) \frac{p_1^2}{2} \right. \\
 &\quad \left. + \frac{1}{8\epsilon} \left| -\frac{1}{2} B_3 \frac{\mu^2 + p_1^2}{E_1 E_2} + A_3 (\uparrow\uparrow) (B_3 + iB_4) \right|^2 \right. \\
 &\quad \left. + \frac{1}{8\epsilon} \left| -\frac{1}{2} B_4 \frac{\mu^2 + p_1^2}{E_1 E_2} + A_4 (\uparrow\uparrow) (B_3 + iB_4) \right|^2 \right. \\
 &\quad \left. + \frac{1}{4\epsilon} \Re \left\{ -\frac{1}{2} B_3 \frac{\mu^2 + p_1^2}{E_1 E_2} + A_3 (\uparrow\uparrow) (B_3 + iB_4) \right\} \right. \\
 &\quad \left. \cdot \left\{ -\frac{1}{2} B_4 \frac{\mu^2 + p_1^2}{E_1 E_2} + A_4 (\uparrow\uparrow) (B_3 + iB_4) \right\} \right]
 \end{aligned}$$

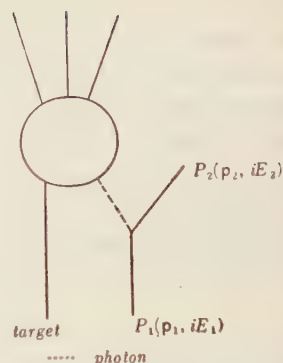


Fig. 1.

$$\begin{aligned}
& + \frac{1}{4\epsilon} \left(\frac{\epsilon\mu}{2E_1E_2} \right)^2 \{ |-B_1 + iB_2|^2 + |B_1 + iB_2|^2 \} \\
& + \frac{1}{2} |B_3 + iB_4|^2 \frac{1}{2\epsilon} \left(\frac{1}{E_1^2} + \frac{1}{E_2^2} \right) \frac{p_{\perp}^2}{2} \Bigg], \quad (12)
\end{aligned}$$

where the expressions of $A_{\nu}(\uparrow\uparrow)$ and $A_4(\uparrow\uparrow)$ are given by (9) in I and $p_{\perp}^2 = p_{1x}^2 + p_{1y}^2 = p_{2x}^2 + p_{2y}^2$. This is exactly the form which we have required. The first term corresponds to the process that the virtual photon polarizes transversely and the spin of the incident particle does not flip.*

§ 4. Derivation of (I) and the condition for applicability of the W-W method

First we shall confine ourselves to the first term in (12). In B_{μ} 's the four-vector k of the virtual photon is contained in the forms

$$k^2 = (\epsilon^2/E_1E_2) (\mu^2 + p_{\perp}^2), \quad (13)$$

and

$$(kq) = \frac{\epsilon q_0}{2E_1E_2} (\mu^2 + p_{\perp}^2) - \frac{\epsilon}{2q_0} (m^2 + q_{\perp}^2), \quad (14)$$

where $q = (\mathbf{q}, iq_0)$ is the four-vector of some particle created in the final state and m its rest mass. To derive (14) we used the small angle approximation (11), $q_3 = q_0 - (1/2q_0) \cdot (m^2 + q_{\perp}^2)$, where $q_{\perp} = |\mathbf{q}| \cdot \alpha$ and α is the zenith angle of \mathbf{q} . From (13) and the factor $1/k^4$ in (12) the integration over p_{\perp}^2 contributes to the integral cross section when its order of magnitude is less than $O(\mu^2)$. Then we can neglect terms containing $(1/E_1E_2) (\mu^2 + p_{\perp}^2)$ in B_{μ} 's as compared with terms like the second term in (14), if

$$\epsilon\mu/E_1m \ll 1. \quad (15)$$

Since the four-vector k satisfied $k^2 = 0$ in B_{μ} 's under the condition (15), the factor $(1/2) \{ |B_1|^2 + |B_2|^2 \} (1/2\epsilon)$ is completely in accordance with the factor in (3) arising from the corresponding process caused by a real photon. Therefore, the contribution of the first term to the cross section becomes from (3) and (6)

$$\int \frac{1}{2} \{ |B_1|^2 + |B_2|^2 \} \frac{1}{2\epsilon} \rho'_{\mathbf{k}q'}(\epsilon) d\epsilon = \int \frac{1}{2} \{ |B_1|^2 + |B_2|^2 \} \frac{1}{2\epsilon} \rho'_{\mathbf{k}q'}(\epsilon) = \sigma_{\tau}(\epsilon) d\epsilon q'(\epsilon),$$

where

* For a particle of spin 0, the same procedure gives

$$\sum_{\varphi_2} |M|^2 = \frac{8\pi\epsilon(Ze)^2}{k^4} \left[\frac{1}{2} \{ |B_1|^2 + |B_2|^2 \} \frac{1}{2\epsilon} \frac{p_{\perp}^2}{E_1E_2} + \frac{1}{8\epsilon} \frac{(E_1 + E_2)^2}{E_1E_2} |B_3 + iB_4|^2 \right].$$

As is readily seen, the first and the second terms above correspond respectively to the first term in (12) and the terms containing $|B_3 + iB_4|^2$ in (12).

$$q'(\epsilon) = 8\pi\epsilon(Ze)^2 \int \frac{1}{k^4} \left(\frac{1}{E_1^2} + \frac{1}{E_2^2} \right) \frac{p_\perp^2}{2} \rho.$$

Expressing ρ in the S-system (see Appendix in I), and from (13), $q'(\epsilon)$ can be rewritten as

$$q'(\epsilon) = \frac{(Ze)^2}{(2\pi)^3} \frac{4\pi}{\epsilon} \int \int \frac{p_\perp^2 d(p_\perp^2) d\varphi_2}{(\mu^2 + p_\perp^2)^2}. \quad (16)$$

Since the denominators of D_F and S_F functions contained in B_μ 's are the sum of the terms like (13) and (14), the contribution of p_\perp^2 to the cross section falls off rapidly when $p_\perp^2 \gtrsim (mE_1/\epsilon)^2$. Thus integrating over p_\perp^2 from zero to $(\alpha mE_1/\epsilon)^2$, where α is a constant of order unity, we obtain

$$q'(\epsilon) = (Ze)^2 \frac{2}{\pi} \frac{1}{\epsilon} \left(\log \frac{\alpha mE_1}{\epsilon\mu} - \frac{1}{2} \right). \quad (17)$$

This is, except for a trivial difference, nothing but the spectrum, (1), of the virtual photons used in the W-W method.

As is easily seen the fifth term in (12), which corresponds to the process in which the spin of the incident particle flips, can be neglected as compared with the first. Therefore, under the condition (15) the contribution from the transverse photon is exactly the same as that derived by the W-W method.

Now we will seek the condition that the contributions from longitudinal and scalar photons can be neglected as compared with that from transverse photons. To begin with, let us compare the orders of magnitude of the two terms, $(1/2)B_3(\mu^2 + p_\perp^2)/(E_1E_2)$ and $A_3(\uparrow\uparrow)(B_3 + iB_4)$. The main term of the latter is $B_3 + iB_4(-F(k_3, \epsilon))$ from (9) in I, where $k_3 = \epsilon[1 + (1/E_1E_2)(\mu^2 + p_\perp^2)]^{1/2}$ is the z -component of the momentum of the virtual photon. We expand $F(k_3, \epsilon)$:

$$F(k_3, \epsilon) = F(\epsilon, \epsilon) + (dF/dk_3)_{k_3=\epsilon} \Delta\epsilon, \\ \Delta\epsilon \cong (\epsilon/2E_1E_2)(\mu^2 + p_\perp^2).$$

The first term $F(\epsilon, \epsilon)$ is exactly equal to $B_3^\dagger + iB_4^\dagger$ which is the corresponding term in the process caused by a real photon and vanishes from gauge invariance. Then the comparison of the relative orders of the above two terms is reduced to the comparison of B_3 with $\epsilon\dot{F}(\epsilon, \epsilon)$.* Therefore we shall consider the relative order of $\dot{F}(\epsilon, \epsilon)$ to B_3 or B_4 . Omitting the spinors of particles in the initial and final states and constant factors, B_μ is generally given by a sum of terms of the type

$$\gamma_\mu S_F \gamma_\alpha S_F \gamma_\beta \dots D_F \dots^{**}. \quad (18)$$

* $\dot{F}(\epsilon, \epsilon)$ means $(dF/dk_3)_{k_3=\epsilon}$, and so forth.

** When the interactions contain derivative couplings or mesons contribute to the process, (18) must be modified but the similar arguments can be developed.

Here γ_μ arises from the vertex part at which the virtual photon interacts and its position in (18) is different from diagram to diagram and from process to process but the following argument does not depend on its position. From (18) $\dot{F}(\epsilon, \epsilon)$ is expressed as

$$\dot{F}(\epsilon, \epsilon) = (\gamma_3 + i\gamma_4) \sum [\dot{S}_F \gamma_\alpha S_F \dots + S_F \gamma_\alpha \dot{S}_F \dots + \dots], \quad (19)$$

Expressing by $\sum p_i$ the vector sum of the four-vectors of the particles in the final state, the forms of S_F and D_F functions are generally

$$S_F = \frac{i\gamma(k + \sum p_i) - m}{(k + \sum p_i)^2 + m^2}, \quad D_F = \frac{1}{(k + \sum p_i)^2},$$

and we obtain

$$\begin{aligned} \dot{S}_F &= \frac{i\dot{\gamma}_3}{(k + \sum p_i)^2 + m^2} - \frac{2(\epsilon + \sum p_{i3})}{(k + \sum p_i)^2 + m^2} S_F, \\ \dot{D}_F &= -\frac{2(\epsilon + \sum p_{i3})}{(k + \sum p_i)^2} D_F. \end{aligned}$$

Since $0(\epsilon + \sum p_{i3}) = 0(\epsilon)$ from the assumption A and $0[(k + \sum p_i)^2 + m^2] = 0[(k + \sum p_i)^2] = 0(m^2)$ from (14), we obtain

$$\dot{S}_F = 0(\epsilon/m^2) S_F, \quad \dot{D}_F = 0(\epsilon/m^2) D_F.$$

That is, the differentiation of S_F and D_F by k_3 changes the order of magnitude relative to S_F and D_F by $0(\epsilon/m^2)$. Then (19) is rewritten as

$$\begin{aligned} \dot{F}(\epsilon, \epsilon) &\cong (\gamma_3 + i\gamma_4) (\epsilon/m^2) \sum [aS_F \gamma_\alpha S_F \dots + bS_F \gamma_\alpha S_F \dots + \dots] \\ &\equiv (\epsilon/m^2) (B_3' + iB_4'), \end{aligned} \quad (20)$$

where a, b, \dots are quantities of order unity and

$$B'_\mu = \gamma_\mu \sum [aS_F \gamma_\alpha S_F \dots + bS_F \gamma_\alpha S_F \dots + \dots].$$

The B'_μ 's are not identical with the B_μ 's and the gauge invariance for $(B_3' + iB_4')_{k_3=\epsilon}$ does not hold as it does for $(B_3 + iB_4)_{k_3=\epsilon}$. But the B'_μ 's are of the same order of magnitude as the B_μ 's. Therefore

$$0|(B_3' + iB_4')/B_3| = 0|(B_3' + iB_4')/B_3'|.$$

Summarizing the above arguments we obtain

$$0 \left| \frac{A_3(\uparrow\uparrow)}{B_3(\mu^2 + p_\perp^2)/E_1 E_2} (B_3 + iB_4) \right| = 0(\epsilon^2/m^2) 0 \left| \frac{B_3' + iB_4'}{B_3'} \right|.$$

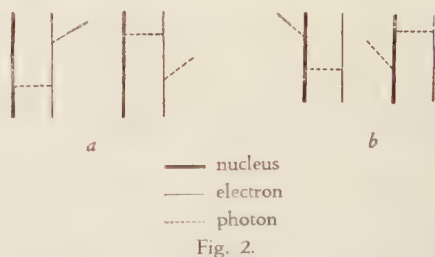
Furthermore the order of $(B_3' + iB_4')/B_3'$ is equal to $0(m^2/\epsilon^2)$ from the result given in Appendix. Thus the two terms contained in the second term in (12) are of the same order. In the same manner we see that all terms from the second to fourth are of the same order and, as is easily seen, the last term may be neglected in comparison with these. Therefore, in order to derive the desired condition it is sufficient to consider only

the relative order of the second and first terms that is, $O(\mu/E_1)O|B|B_1|$. From the result given in Appendix this order is $O(\epsilon\mu/E_1m)$. This fact shows simply that (15) is also the condition that the contributions from the longitudinal and scalar photons can be neglected in comparison with that from the transverse photons. That is, the W-W method is rigorously applicable under the condition (15).

§ 5. Discussions on our assumption and concluding remarks

Until now, we have neglected the diagrams except for the one of the type shown in Fig. 1. This neglect is guaranteed in cases that created particles in the final state are due to nuclear interactions between these particles and the target, e.g., the π production by μ -meson, since the lowest order diagrams with respect to the electromagnetic interaction of an incident particle are reduced to only one type given in Fig. 1. However, in pure electromagnetic phenomena it is necessary to assume that the mass of the fast particle is heavier than that of an electron. Let us consider,

for example, Bremsstrahlung by electron. There are two types of diagrams given in Fig. 2a and 2b, the former of which are considered as Compton scattering of a virtual photon by the electron and the latter as that of a virtual photon by the nucleus. Since the cross section of Compton scattering is inversely proportional



to the square of the mass of the scatterer, the contribution from the diagrams 2b can be neglected as compared with that from 2a. The situation is also the same for the pair creation by μ -meson. Therefore, in pure electromagnetic processes the diagrams of type 2b can be always neglected, provided that the mass of the incident particle is large compared with that of an electron. In other words we can neglect the diagrams that the incident particle interacts twice or more with the photon field. In cases in which the incident particle is an electron, it will be necessary to take into account the diagrams not considered in our treatment, though for the pair creation by an electron we have shown that the contributions from the diagrams of type 2b can be neglected⁽⁶⁾.

Next we shall discuss on the assumption A. When a target particle is very heavy, for example a nucleus, it can be considered to be at rest during the collision and its action on the incident particle can be replaced by an external field. A pair creation by a charged particle is this case. The reader will see that such a case is included in our treatment. Furthermore, we can see that the condition (15) holds not only for such a case, but also for cases where some of particles created in the final state have non-relativistic energies. Because the main terms of B_i arise from P_j 's of the particles having relativistic energies, not from those having non-relativistic energies, and the argument given in Appendix is valid for such main terms.

The W-W method gives the same results as the F-D method as long as the condition (15) holds. On the other hand it must be noted that the main contribution to

the cross section arises from the region satisfying (15), i.e., the region where the transferred energy is small. For $\epsilon/\mu/E_1 m > 1$ the terms containing the factor $(1/E_1 E_2)(\mu^2 + p_\perp^2)$ can not be neglected both in the numerators and the denominators in B_μ . The increase in the denominators due to addition of these terms is larger than those in the numerators, as shown by the dimensional analysis of matrix elements. Furthermore the factor $1/k^4$ in (12) suppresses the integral value over p_\perp^2 more largely in such region than in the region satisfying (15). Therefore, the cross section derived by the W-W method will give the correct order of magnitude even if we take into account the contribution from the region where the condition (15) does not hold. This fact was explicitly shown in the case of pair creation by charged particle in I. The condition, (15), gives a rough measure of the most probable value of the transferred energy. The heavier the mass of the created particle, the more extensive is the region which gives the main contribution. For example, when a μ -meson is incident, the rate of transferred energy ϵ/E_1 in that region lies between zero and nearly $1/200$ for the pair creation, but between zero and nearly one for π -production.

In conclusion the authors should like to express their gratitude to Prof. T. Inoue for his encouragement and also Drs. H. Tanaka and H. Hasegawa for their advice and discussions.

Appendix

Since B_μ is a component of a tensor of rank one, $\sum_{\text{spin}} |B_\mu|^2$ (not summed over μ) becomes that of rank two and is expressed as a sum of tensors of four types, i.e.,

$$\sum_{\text{spin}} |B_\mu|^2 = \sum_{\text{spin}} \sum_{\gamma\gamma ee'} \{ \delta_{\mu\mu} (\dot{p}q) (ee') + \dot{p}_\mu q_\mu (ee') + \dot{p}_\mu e_\mu (qe') + e_\mu e'_\mu (\dot{p}q) \} \cdot C^{\gamma\gamma ee'}(r, s, \dots, e'', e''', \dots), \quad (\text{A1})$$

where \dot{p}, q, r, \dots are four vectors of particles in the final state or that of the virtual photon and e, e', e'', \dots are polarization vectors of photons in the final state if emitted.* $C^{\gamma\gamma ee'}(r, s, \dots, e'', e''', \dots)$ consist of the scalar products of $r, s, \dots, e'', e''', \dots$ except \dot{p}, q, e and e' and are independent of the index μ . $\sum_{\gamma\gamma ee'}$ means the summation over all the combinations of two energy-momentum four-vectors and two polarization vectors extracted from vectors involved in B_μ .

Now what we want to know is how the order of magnitude of $\sum_{\text{spin}} |B_\mu|^2$ changes when we replace $\mu=1$ or 2 by 3 or 4. This is nothing but to know the relative order of each term contained in the braces in (A1) by changing the index μ . Using the assumption A and the fact that the order of magnitude of the transverse momentum of each particle is $O(m)$, we obtain the relative order of magnitude of each term in (A1) as given in the table 1, where the order of the magnitude of the polarization vector is

* If no photon be emitted, (A1) becomes

$$\sum_{\text{spin}} |B_\mu|^2 = \sum_{\text{spin}} \sum_{\gamma\gamma} \{ \delta_{\mu\mu} (pq) + \dot{p}_\mu q_\mu \} C^{\gamma\gamma}(r, s, \dots),$$

and our following arguments are not changed.

taken as $0(e_1)=0(e_2)=0(1)$ and $0(e_3)=0(m/\epsilon)$. Therefore considering that the main terms of longitudinal and scalar parts are $0(\epsilon^2)$ and that of transverse part $0(m^2)$, we get

$$0|B_{3,4}/B_{1,2}|^2=0(\epsilon^2/m^2). \tag{A2}$$

In a similar manner we obtain the relative order of $B_3'+iB_4'$ by replacing γ_μ by $\gamma_3+i\gamma_4$ in the main terms of B_3 , thus giving table 2. Hence we get

$$0|(B_3'+iB_4')/B_3'|^2=0(m^4/\epsilon^4). \tag{A3}$$

Table 1.

| μ | 1 | 2 | 3 | 4 |
|----------------------------|----------|----------|---------------------|-----------------|
| $\delta_{\mu\mu}(pq)(ee')$ | $0(m^2)$ | $0(m^2)$ | $0(m^2)$ | $0(m^2)$ |
| $p_\mu q_\mu(ee')$ | $0(m^2)$ | $0(m^2)$ | $0(\epsilon^2)$ | $0(\epsilon^2)$ |
| $p_\mu e_\mu(qe')$ | $0(m^2)$ | $0(m^2)$ | $0(m^2)$ | |
| $e_\mu e_\mu'(pq)$ | $0(m^2)$ | $0(m^2)$ | $0(m^4/\epsilon^2)$ | |

Table 2.

| | $\gamma_3+i\gamma_4$ | γ_3 |
|----------------------------|--|--------------------------------|
| $\delta_{\mu\mu}(pq)(ee')$ | $(pq)(ee')(1^2+i^2)=0$ | 1 $0(m^2)$ |
| $p_\mu q_\mu(ee')$ | $(p_3+i p_4)(q_3+i q_4)(ee')$ $0(m^4/\epsilon^2)$ | $p_3 q_3$ $0(\epsilon^2)$ |
| $p_\mu e_\mu(qe')$ | $(e_3+i e_4)(p_3+i p_4)(qe)$ $0(m^4/\epsilon^2)$ | $e_3 p_3$ $0(m^2)$ |
| $e_\mu e_\mu'(pq)$ | $(e_3+i e_4)(e_3'+i e_4')(pq)$ $0(m^4/\epsilon^2)$ | $e_3 e_3'$ $0(m^4/\epsilon^2)$ |

References

1) E. J. Williams, Kgl. Dansk. Vid. Selsk. 13 (1935), No. 4.
2) C. F. v. Weizsäcker, Zeit. f. Phys. 88 (1934), 612.
3) E. J. Williams, loc. cit..
4) Y. Nishina, S. Tomonaga and M. Kobayashi, Sci. Pap. Inst. Phys. Chem. Research, Japan, 27 (1935), 137.
E. J. Williams, loc. cit.
H. J. Bhabha, Proc. Roy. Soc. A 152 (1935), 559.
5) H. Fukuda, Y. Fujimoto and M. Keshiba, Prog. Theor. Phys. 6 (1951), 788.
6) T. Murota, A. Ueda and H. Tanaka, Prog. Theor. Phys. 16 (1956) 482, quoted as I here.

Helium Capturing Reactions in Stars

Satio HAYAKAWA*, Chushiro HAYASHI**, Mitsuo IMOTO**
and Ken KIKUCHI*** *)

* *Research Institute for Fundamental Physics, Kyoto University Kyoto*

** *Department of Physics, Kyoto University, Kyoto*

*** *Department of Physics, Osaka University, Osaka*

(Received July 12, 1956)

The generation of ^{12}C , ^{16}O and ^{20}Ne by successive α -capturing processes in the helium core of stars is calculated under the condition at temperature $\sim 10^8\text{K}$. The rates of forming ^{12}C and ^{20}Ne are found to be much larger than the previous calculations by Salpeter, Öpik and Hoyle, due to the resonance levels recently discovered. The production rates of ^{24}Mg to ^{40}Ca are also estimated. The formation of these nuclei is found to be negligible at temperatures below $4 \times 10^8\text{K}$. The luminosity of such stars of globular clusters belonging to the horizontal branch in the H-R diagram that are supposed to be burning helium is accounted for in terms of the above thermonuclear reactions at the central temperature of $(1.40 \pm 0.10) \times 10^8\text{K}$ and the density of about 10^3g cm^{-3} . The lifetime of a star in this evolutionary stage is estimated roughly as 10^7 years. In this course of evolution ^{20}Ne is formed more abundant than others, in disagreement with the average cosmic abundances of ^{12}C , ^{16}O and ^{20}Ne . The modification of the abundances after this stage is suggested as necessary.

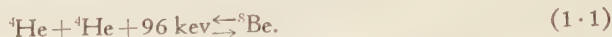
§ 1. Introduction and summary

The energy generation in stars belonging to the main sequence has been accounted for in terms of the "carbon-nitrogen cycle"^{1),3)} and the "proton-proton chain"²⁾. Owing to these nuclear reactions hydrogen is converted into helium, so that the concentration of helium in the cores of these stars increases, as they evolve. The helium nuclei thus produced can hardly be converted into heavier nuclei by the reactions involving light nuclei alone, because the nuclei of mass numbers 5 and 8 are unstable and even stable isotopes of Li, Be and B return rapidly back to ^3He and ^4He due to the absorption of protons which are rich in the core of such stars³⁾. However, the synthesis of carbon and heavier nuclei is regarded as necessary in view of the evolution scheme, as emphasized by a number of authors.⁴⁾ Some stars belonging to the "giant branch" are believed to undergo such a nuclear synthesis that ^4He nuclei are converted into ^{12}C and heavier ones. The synthesis of ^{12}C seems to be required also to explain the over-abundance of ^{12}C observed on the surfaces of R, N and Wolf-Rayet stars.

A possibility of such a synthesis process has been suggested by Salpeter⁵⁾ as to take place in hydrogen exhausted stars, when their central temperature becomes higher than

*) Now at Institute for Nuclear Study, Tokyo University

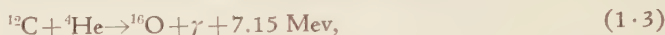
10^8 K. At such a high temperature an appreciable amount of ${}^8\text{Be}$ nuclei is formed under the thermal equilibrium with ${}^4\text{He}$ nuclei:



A ${}^8\text{Be}$ nucleus thus formed can be converted into ${}^{12}\text{C}$ by capturing a ${}^4\text{He}$ nucleus:



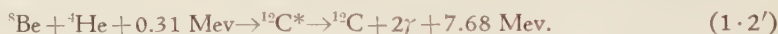
These processes are supposed to supply a sufficient amount of energy for about 10^7 years and may explain the existence of the ${}^{12}\text{C}$ -rich stars. He has further suggested the synthesis of heavier nuclei as due to successive α -capturing reactions:



and



In those days the knowledge about the properties of these nuclei was rather poor, so that his arguments were obliged to be qualitative. Since a level of ${}^{12}\text{C}$ at 7.68 Mev and some properties of other nuclei were observed, Hoyle⁶ tried to calculate the reaction rates of (1.2) and (1.3), taking into account the resonance capture. But the lack in knowledge of the spins and parities of the levels forbade him to deduce reliable reaction rates, so that he rather predicted them in reference to the mean cosmic abundances of ${}^{12}\text{C}$ and ${}^{16}\text{O}$. Lately Salpeter⁷ noticed the role of the 7.68 Mev level, which was supposed to have zero spin and even parity, and proposed a process,



But its details and results are not reported yet.

In order to obtain a quantitative answer of the synthesis of heavy nuclei, we attempted to calculate the reaction rates of (1.1) to (1.4), on the basis of recent nuclear data.⁸⁾ We also made a rough estimate on the capture of ${}^4\text{He}$ by ${}^{20}\text{Ne}$ to ${}^{36}\text{Ar}$. Based on our calculated reaction rates and the rates of energy generation, we briefly discussed the synthesis of elements as well as the evolution of stars.

In § 2 the general method of treating the thermonuclear reaction is presented, taking account of the contributions from resonant as well as non-resonant processes. For a resonance energy near the Gamow peak, the conventional method of the Gamow peak is modified, while, for the sufficient separation between the resonance and Gamow peaks, the contributions from these two are carefully examined. Our formulas are not qualitatively different from conventional ones, but are useful for quantitative studies. In § 3 the formation of ${}^8\text{Be}$, and the rates of formation of ${}^{12}\text{C}$, ${}^{16}\text{O}$ and ${}^{20}\text{Ne}$ and also the rates of the energy generation are calculated. The result is summarized in Table 3 and in Fig. 5. In § 4 the formation of ${}^{24}\text{Mg}$ to ${}^{40}\text{Ca}$ is discussed. This is found to be negligible at temperatures below 4×10^8 K. In § 5 the luminosity of such stars that undergo the helium burning processes is interpreted in terms of a polytrope model of stars. For the stars of globular clusters in the horizontal branch in the H-R diagram, the central temperature and density are found to be about 1.4×10^8 K and 10^3 g cm^{-3} . The reason why these figures are much smaller than expected by Salpeter⁵⁾ is due

mainly to the resonance level of ^{12}C . The relative abundances of ^{12}C , ^{16}O and ^{20}Ne are calculated for given temperatures. In contrast to Hoyle⁽⁶⁾ the abundance of ^{16}O is much lower than that of ^{20}Ne under the above conditions. This is due to the newly discovered level⁽¹⁵⁾ of ^{20}Ne . The modification of the abundance after this evolutionary stage is suggested as necessary. The mean lifetime of the helium burning stars is estimated as about 10^7 years under the same condition and this is in agreement with the observed abundance of the stars in the giant branch. Our result is in essential agreement with that of Obi et al.⁽⁸⁾ In § 6 our conclusions are summarized and discussions are made on related problems.

§ 2. General formulas

Let us assume a gas composed of two kinds of nuclei 1 and 2, each of which is uniform in space and obeys the Maxwellian distribution of velocities at given temperature T . Suppose that these nuclei of respective densities, n_1 and n_2 , collide with one another with reaction cross section $\sigma(E)$, where E is the kinetic energy in the center of mass system of the colliding nuclei. Then the reaction rate per unit time and per unit volume is given by

$$P = \frac{4n_1n_2}{(2\pi\mu)^{1/2}(kT)^{3/2}} \int_0^\infty \sigma(E) \exp(-E/kT) E dE, \quad (2.1)$$

where k is the Boltzmann constant. μ is the reduced mass and is expressed, in terms of the masses of the respective nuclei, m_1 and m_2 , as $\mu = m_1m_2/(m_1+m_2)$.

The reaction is regarded as to form a compound nucleus. The energy level of the compound nucleus is measured from the sum of mass energies of the colliding nuclei and is designated as E_n . The property of a level is implied in the widths for the absorption and the subsequent emission, Γ_{sn} and Γ_{in} , respectively. S_1 and S_2 show the spins of target and incident particles, respectively. By employing these quantities the cross section is expressed as

$$\sigma(E) = \frac{1}{(2S_1+1)(2S_2+1)} \frac{\pi\hbar^2}{2\mu E} \sum_n \frac{(2J_n+1)\Gamma_{sn}\Gamma_{in}}{(E-E_n)^2 + \Gamma_n^2/4}. \quad (2.2)$$

The summation is extended over all possible levels corresponding to spin J_n , which can consist of any particular orbital angular momentum l_n , but one needs in most cases to take only one level into account. The total width Γ_n is practically the sum of Γ_{sn} and Γ_{in} .

The absorption width Γ_s can be expressed by a classical formula as

$$\Gamma_s = G_s \exp(-2C_l), \quad (2.3)$$

with

$$C_l = -2\sqrt{2\mu Z_1 Z_2 e^2 R/\hbar^2} + \pi Z_1 Z_2 e^2 \sqrt{\mu/\hbar} \sqrt{2E} + l(l+1)\hbar/(2Z_1 Z_2 e^2 R\mu)^{1/2}, \quad (2.4)$$

and

$$G_s = 2D_s/\pi \cdot \sqrt{(B-V)/(V+E)} \cong 2D_s/\pi \cdot \sqrt{B/V}. \quad (2.5)$$

Z_1e and Z_2e are the charges of the respective nuclei and B the height of the Coulomb barrier between them. D_s is the mean level distance for the levels with the same spin and parity as the absorbing level and V the depth of the potential for one particle against the other. These two quantities are rather uncertain, but their influence on the final result is relatively unimportant. We evaluate D_s from observed levels and choose V for the helium capturing process as

$$V = K_0^2 \hbar^2 / 2M_\alpha \simeq 5.1 \text{ Mev}, \quad (2.6)$$

where $K_0 = 1 \times 10^{-13} \text{ cm}^{-1}$ is supposed to be the wave number of an α -particle just inside the nuclear surface.⁹⁾ The largest uncertainty in the cross section arises from the channel radius R . In reference to various induced reactions we use the expression

$$R = R_1 + R_\alpha, \quad (2.7)$$

with uncertainty of $0.5 \times 10^{-13} \text{ cm}$. $R_\alpha = 1.6 \times 10^{-13} \text{ cm}$ is chosen from the α - α scattering and R_1 is $2.4 \times 10^{-13} \text{ cm}$ for ^8Be , $3.8 \times 10^{-13} \text{ cm}$ for ^{12}C , $4.3 \times 10^{-13} \text{ cm}$ for ^{16}O and $1.3 A_1^{1/3} \times 10^{-13} \text{ cm}$ for ^{20}Ne and heavier nuclei of mass number A_1 .

Among various quantities introduced above, E_n and l'_n are regarded as independent of energy, because the energy interval under consideration is relatively narrow. Taking such energy insensitive quantities out of the integral, (2.1) is reduced to

$$P = \frac{\sqrt{2\pi} \hbar^2 n_1 n_2}{(\mu k T)^{3/2} (2S_1 + 1) (2S_2 + 1)} \sum_n (2J_n + 1) \Gamma_n G_n \exp \left[\left(4 \sqrt{\frac{2\mu Z_1 Z_2 e^2 R}{\hbar^2}} - 2l(l+1)\hbar / (2Z_1 Z_2 e^2 R \mu)^{1/2} \right) \times \int_0^\infty \frac{\exp(-b/\sqrt{E - E_n})}{(E - E_n)^2 + l_n'^2/4} dE \right], \quad (2.8)$$

where

$$b = 2\pi Z_1 Z_2 / \hbar \cdot \sqrt{\mu/2} = 1.00 Z_1 Z_2 \sqrt{A_1 A_2 / (A_1 + A_2)} \text{ Mev}^{1/2} \\ = 1.27 \times 10^{-3} Z_1 Z_2 \sqrt{A_1 A_2 / (A_1 + A_2)} \text{ erg}^{1/2}. \quad (2.9)$$

The numerator in the integrand has the so-called Gamow peak at

$$E_g = (bkT/2)^{2/3} = 0.0264 [A_1 A_2 Z_1^2 Z_2^2 / (A_1 + A_2)]^{1/6} T^{5/6} \text{ Mev}, \quad (2.10)$$

and the width of the peak is

$$\gamma_g = (4E_g kT/3)^{1/2} = 0.0178 [A_1 A_2 Z_1^2 Z_2^2 / (A_1 + A_2)]^{1/6} T^{5/6} \text{ Mev}. \quad (2.11)$$

Here and hereafter T is measured in 10^8 K and all energies are in Mev, if not specially mentioned.

In the presence of the resonance denominator in (2.8), the integrand may have another peak at about $E = E_n$. In the integration of (2.8), therefore, we have to take at least two peaks into account. One comes from the peak similar to the Gamow peak, but is modified slightly due to the denominator. The other is the resonance peak that plays a dominant role for $|E_n - E_n| \cdot kT \lesssim \gamma_g$. The approximate methods used in respective cases are described in what follows.

Modified Gamow peak or non-resonance peak. A peak appears at E_0 which is close to E_g and its width is γ_0 . E_0 and γ_0 are given by

$$\frac{bkT}{2E_0^{3/2}} = \frac{2(E_0 - E_n)kT + (E_0 - E_n)^2 + \Gamma_n^2/4}{(E_0 - E_n)^2 + \Gamma_n^2/4} \quad (2.12a)$$

and

$$\frac{1}{\gamma_0^2} = \frac{3}{4E_0kT} - \frac{2(E_0 - E_n)^2 - \Gamma_n^2/2}{\{(E_0 - E_n)^2 + \Gamma_n^2/4\}^2}. \quad (2.12b)$$

Then the integral is evaluated as

$$I_n = \int_0^\infty \frac{\exp(-b/\sqrt{E - E_0}/kT)}{(E - E_n)^2 + \Gamma_n^2/4} dE \simeq \frac{\exp(-b/\sqrt{E_0 - E_0}/kT)}{(E_0 - E_n)^2 + \Gamma_n^2/4} \int_0^\infty \exp[-(E - E_0)^2/\gamma_0^2] dE,$$

or

$$I_n = \sqrt{\pi} \gamma_0 \frac{\exp(-b/\sqrt{E_0 - E_0}/kT)}{(E_0 - E_n)^2 + \Gamma_n^2/4}. \quad (2.13)$$

It is readily seen that E_0 and γ_0 approach E_g and γ_g respectively for $|E_g - E_n| > \gamma_g \gtrsim kT \gg \Gamma_n$. In this case the Gamow approximation is subject to the following errors:

$$E_0 \simeq \left\{ 1 - \frac{4kT}{3(E_g - E_n)} \right\} E_g, \quad (2.12a')$$

and

$$\gamma_0 \simeq \left\{ 1 + \frac{2kT}{3(E_g - E_n)} \right\} \gamma_g. \quad (2.12b')$$

The contribution from distant levels is of the order of $(E_g - E_n)^2/D_s^2$ in comparison with that from the nearest level at E_n . If the contribution from the non-resonance peak is important, $|E_g - E_n|$ is far smaller than D_s , so that the contribution from distant levels is negligible.

Resonance peak. The integration in this case can be evaluated as

$$\begin{aligned} I_r &\simeq \exp(-b/\sqrt{E_n - E_n}/kT) \int_0^\infty [(E - E_n)^2 + \Gamma_n^2/4]^{-1} dE \\ &= \left(\frac{2}{\Gamma_n} \right) \left[\frac{\pi}{2} - \tan^{-1} \left(\frac{2E_n}{\Gamma_n} \right) \right] \exp \left(\frac{-b}{\sqrt{E_n}} - \frac{E_n}{kT} \right). \end{aligned}$$

Since $E_n > \Gamma_n$ in most cases, this can be approximated as

$$I_r = (2\pi/\Gamma_n) \exp(-b/\sqrt{E_n - E_n}/kT). \quad (2.14)$$

The contribution from distant levels is also negligible, because the exponential factor becomes extremely small.

In order to compare the contributions from the resonance and the non-resonance peaks, we consider three cases, according to the magnitude of $|E_0 - E_n|$.

(a) $|E_0 - E_n| < \Gamma_n$, γ_0 . Two peaks overlap with one another and $\gamma_0 \simeq \Gamma_n/2$. Hence either of (2.13) or (2.14) gives us essentially the same result.

(b) $\Gamma_n < |E_0 - E_n| \lesssim \gamma_0$. There is a sharp resonance peak in a broad Gamow peak.

Since $\gamma_0 \gtrsim |E_0 - E_n|$ and the exponential factors in (2.13) and (2.14) are of the same order of magnitude, we have $I_n/I_r \simeq I'_n \gamma_0 / |E_0 - E_n|^2$. Hence the resonance peak represents an essential part of the integral I .

(c) $I'_n \gamma_0 < |E_0 - E_n|$. Two peaks are separated from one another. The contribution from the non-resonance peak approaches the Gamow approximation, as E_n becomes far from E_0 , so that $E_0 \simeq E_g$ and $\gamma_0 \simeq \gamma_g$. Comparison has to be made with the numerical values of I'_n , γ_0 and $|E_0 - E_n|$. With increasing $|E_0 - E_n|$, I_n becomes more important than I_r , because the exponential factor in I_r decreases more rapidly than $(E_0 - E_n)^{-2}$ in I_n .

In the resonance reaction the following approximation is worth remarking. There often occurs either $I'_n \gg I'_{sn}$ or $I'_n \ll I'_{sn}$. Consequently, one may set either $I'_n = I'_{sn}$ or $I'_n = I'_{in}$. Since the reaction rate is proportional to $I'_{sn} I'_{in} / I'_n$, only I'_{in} or I'_{sn} remains in either of these cases. In the former case the ambiguity due to the α -width is eliminated, while in the latter case one need not know the radiation width.

Keeping the above remarks in mind, we express the reaction rate as

$$P = \frac{\sqrt{2\pi} \hbar^2 n_1 n_2}{(2S_1 + 1)(2S_2 + 1)(\mu k T)^{3/2}} \times \left[\sum_n \frac{(2J_n + 1) \sqrt{I'_{in} I'_{sn}(E_n) \gamma_0}}{(E_0 - E_n)^2 + I_n'^2/4} e^{-E_0/kT} + \sum_n \frac{(2J_n + 1) 2\pi I'_{in} I'_{sn}(E_n)}{I'_n(E_n)} \right] \times e^{-E_n/kT}. \quad (2.15)$$

The first summation is due to the non-resonance absorption, while the second one due to the resonance absorption. In almost all cases one has to maintain only one term in (2.15). For the nonresonance absorption

$$P_u = \frac{2.30 \times 10^{-12}}{(2S_1 + 1)(2S_2 + 1)} \left(\frac{A_1 + A_2}{A_1 A_2} \right)^{3/2} \sum_n (2J_n + 1) \frac{I'_{in} I'_{sn}(E_n) \gamma_0}{(E_0 - E_n)^2} T^{-3/2} 10^{-50.401 E_0/kT} \text{ cm}^{-3} \text{ sec}^{-1} \quad (2.16)$$

and for the resonance absorption

$$P_r = \frac{8.16 \times 10^{-12}}{(2S_1 + 1)(2S_2 + 1)} \left(\frac{A_1 + A_2}{A_1 A_2} \right)^{3/2} \sum_n (2J_n + 1) \frac{I'_{in} I'_{sn}(E_n)}{I'_n(E_n)} T^{-3/2} 10^{-50.401 E_0/kT} \text{ cm}^{-3} \text{ sec}^{-1} \quad (2.16')$$

where T is measured in unit of 10^{83} K . In each part we usually have to leave only one term.

To this we have to add the effect of screening. According to Salpeter,¹⁰⁾ the weak screening is a good approximation in our case. Then the reaction rate must be multiplied by $\exp(-U_0/kT)$ with

$$-U_0/kT = 1.88 \times 10^{-4} Z_1 Z_2 (Z^2 + Z)^{1/2} A^{-1/2} \rho^{1/2} T^{-3/2}, \quad (2.17)$$

where ρ is the density in g cm^{-3} . Z and A are the atomic number and atomic weight of the main constituent, which is ${}^4\text{He}$ in most of our cases. As we are mainly concerned

with the case in which $\rho \simeq 10^3 \text{ g cm}^{-3}$ and $T \simeq 10^8 \text{ K}$, $-U_0/kT$ is of the order of 10^{-1} . Therefore, the correction due to the screening is of the order of 10% and will not exceed 30%. Since this order of magnitude is implied in the uncertainty in various quantities as well as in our method of approximation, we will not include this correction in our numerical results. As soon as a more accurate calculation becomes possible, the correction will easily be taken into account.

The rate of energy generation is obtained by multiplying the Q value, Q , by the reaction rate and by dividing it by density ρ , as it is usually expressed in unit of $\text{erg g}^{-1} \text{ sec}^{-1}$:

$$\varepsilon \equiv QP/\rho. \quad (2.18)$$

§ 3. Synthesis of ^{12}C , ^{16}O and ^{20}Ne

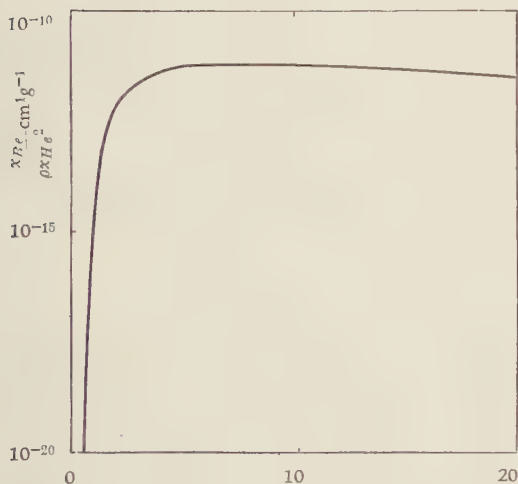


Fig. 1.

T in unit of $10^8 \text{ }^\circ\text{K}$

Concentration of ^8Be in equilibrium with ^4He $x_{8\text{Be}}/\rho x_{\text{He}}^2$ is plotted against T in unit of $10^8 \text{ }^\circ\text{K}$ according to the formula (3.2). $x_{8\text{Be}}$ increases rapidly as temperature increases to $3 \times 10^8 \text{ K}$, but is nearly constant above $4 \times 10^8 \text{ K}$.

As pointed out by Salpeter,⁵⁾ an appreciable amount of ^8Be can be formed from ^4He in the core of a star, in which both temperature and density are very high. The concentration of ^8Be is of primary importance in a series of the helium capturing reactions. This together with the formation of ^{12}C essentially determines the rates of forming subsequent nuclei, ^{16}O and ^{20}Ne .

Since an extremely small fraction of ^8Be nuclei formed in α - α reaction undergoes further α -capture instead of disintegrating back into 2α under the conditions under consideration, the detailed balancing is realized to a good approximation in the reactions $2\alpha \rightleftharpoons \text{Be}^8$. Then, the concentration of ^8Be , $n_{8\text{Be}}$, can be expressed in terms of that of ^4He ,

$n_{8\text{Be}}$ on the basis of the statistical equilibrium as^{*)}

$$\begin{aligned} n_{8\text{Be}}/(n_{\text{He}})^2 &= (2\pi\hbar^2/kT)^{3/2} m_{8\text{Be}}^{3/2}/m_{\text{He}}^3 \cdot \exp(-\varepsilon_0/kT) \\ &= 1.85 \times 10^{-33-4.84/T} \times T^{-3/2} \text{ cm}^3, \end{aligned} \quad (3.1)$$

where $m_{8\text{Be}}$ and m_{He} are the masses of ^8Be and ^4He respectively, and $\varepsilon_0 = 0.096 \text{ Mev}$. is the Q value for the disintegration of ^8Be .

*) $n_{8\text{Be}}/(n_{\text{He}})^2$ given by Hoyle⁶⁾ is smaller than (3.1) by factor two.

Expressing the concentration in terms of the relative concentration defined by

$$\kappa = nm/\rho,$$

we have

$$\kappa_{He}/\kappa_{He}^2 = 5.62 \times 10^{-10-4.84/T} \rho T^{-3/2}. \tag{3.2}$$

This is plotted against T in Fig. 1. κ_{He} increases rapidly as temperature increases to 3×10^8 K. Above 4×10^8 K $\kappa_{He}/(\kappa_{He})^2$ is nearly constant. At $T=10^8$ K and $\rho=10^3$ g cm⁻³, Be⁸ with concentration κ_{Be} of about 10^{-11} can be formed from the assembly of ⁴He. If the capture of an α -particle by ⁸Be is fast enough, an appreciable quantity of ¹²C can be generated, as is seen in the following section. The steep dependence on temperature explains why the synthesis of heavier nuclei via ⁸Be is negligible in other stars of lower temperature, for example in the main sequence stars.

The syntheses of ¹²C, ¹⁶O and ²⁰Ne from ⁸Be thus formed can be discussed separately because the formation of ²⁴Mg becomes comparable only above 4×10^8 K, as shown in § 4. The process which takes part therein is the radiative capture of an α -particle: the particle emission is energetically impossible and the formation of an electron pair is less probable. Hence our first task is to estimate the radiation and α -capturing widths, Γ_i and Γ_s , respectively. For these reactions the knowledge about energy levels can be obtained from existing experiments.¹¹⁾

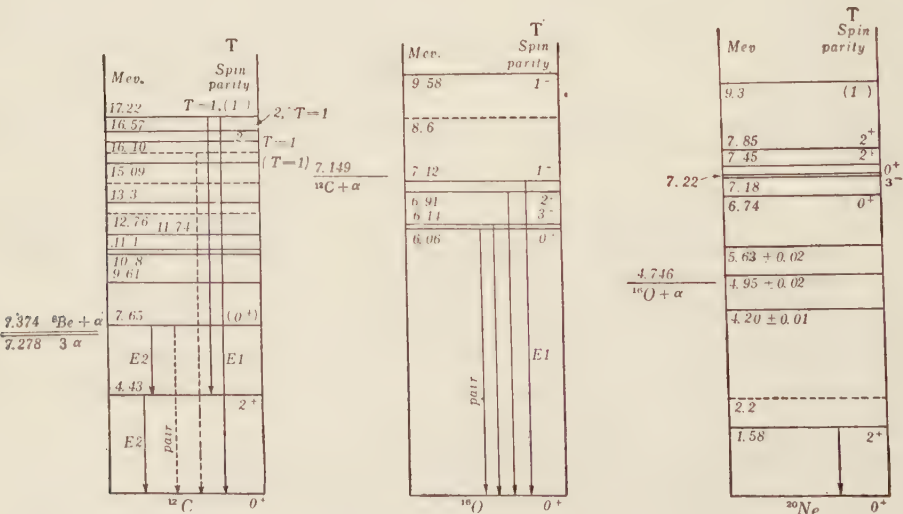
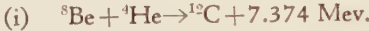


Fig. 2. Level scheme of ¹²C Fig. 3. Level scheme of ¹⁶O Fig. 4. Level scheme of ²⁰Ne

The level scheme of ¹²C is shown in Fig. 2. It is by all means certain that the 7.65 Mev level is only one which contributes to the absorption of an α -particle. The spin and parity of this level are assigned as 0⁺,¹²⁾ so that an S wave α -particle is captured. In calcula-

ting the reaction rate we are allowed to use the resonance formula, based on the following. For $0.7 \lesssim T \lesssim 30$ the contribution from resonance peak is larger than that from the Gamow peak. At $T=21 E_g$ lies in between 7.65 Mev and 9.61 Mev levels. At the resonance energy, $E_n=0.28$ Mev, the α -width $P_s(E_n)$ is found to be about 1 ev, more than three order larger than the radiation width. Consequently, $P_n \simeq P_s$ and P_s/P_n disappears in the formula of the resonance absorption. Thus we have only to evaluate P_i .

Since the spins and parities of both the absorbing level and the ground state are 0^+ , the transition to the ground state by emitting one photon is forbidden. The radiative transition to the ground state can take place by emitting two or more photons. It is usual that the internal pair formation is more probable than the two photon emission in a zero-zero transition. In competing this process, the radiative cascade transition through the 4.43 Mev level with 2^+ has to be taken into account. In the last case only the width for the first E2 transition plays a role in the reaction rate. The widths for these processes are calculated in the following way.

In the conventional formula for such widths the nuclear matrix element is expressed in terms of a "nuclear radius", $A^{1/3}R_0$.⁹⁾ For the $0^+ - 0^+$ transition the matrix element is estimated by Schiff.¹³⁾ In reference to this we estimate $R_0=1.1 \times 10^{-13}$ cm. For the two photon transition we take the 17.22 Mev level with 1^- and isotopic spin unity as an intermediate state and adopt the same value of R_0 . For the E2 transition the nuclear matrix element is different and so is R_0 . The analysis of excitation to the first excited state by high energy electrons results in the E2 width of 10^{-2} ev for the transition from the first to ground states.¹¹⁾¹⁴⁾ Taking only the energy dependence into account the E2 width under consideration is predicted as about 10^{-3} ev. Since the overlap of nuclear states is considered as poorer in the latter case than in the former case, this may be an over-estimate. It may be a reasonable estimate to choose $R_0=1.3 \times 10^{-13}$ cm. Consequently the E2 width is estimated as 7×10^{-4} ev. These results are summarized in Table 1.

Table 1. Radiation and electron emission for ^{12}C

| Process | Transition energy (Mev) | $R_0 (10^{-13} \text{ cm})$ | Width (Mev) |
|--------------------|-------------------------|-----------------------------|---|
| E 2 | 7.65 \rightarrow 4.43 | 1.3 | $\Gamma_{E2} = (7 \pm 3) \times 10^{-10}$ |
| Internal pair | 7.65 \rightarrow 0 | 1.1 | $\Gamma_{pair} = 3 \times 10^{-11}$ |
| 2γ emission | 7.65 \rightarrow 0 | 1.1 | $\Gamma_{2\gamma} = 5 \times 10^{-12}$ |

From Table 1 we conclude that *only the E2 transition needs to be taken into consideration*. On account of the uncertainty in Γ_{E2} , we shall leave Γ_{E2} in the final formula.

Now we are able to give the reaction rate for $0.7 \lesssim T \lesssim 20$. This is presented at an arbitrary value of T in this interval.

$$P_{Be + \alpha \rightarrow C} = \left(\frac{\Gamma_{E2}}{7 \times 10^{-10}} \right) \frac{6.08 \times 10^{-22} - 13.91/T}{T^{3/2}} \text{ cm}^3 \text{ sec}^{-1}. \quad (3.3)$$

This together with (3.1) leads us to

$$P(3\ ^4\text{He} \rightarrow\ ^{12}\text{C}) = p(^8\text{Be} +\ ^4\text{He} \rightarrow\ ^{12}\text{C}) n_{\text{He}} n_{\text{Be}} = p_{3\alpha \rightarrow c} (n_{\text{He}})^3, \quad (3.4)$$

with

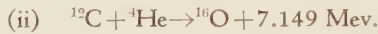
$$p_{3\alpha \rightarrow c} = \left(\frac{\Gamma_{E2}}{7 \times 10^{-10}} \right) \frac{1.12 \times 10^{-54-18.75/T}}{T^3} \text{ cm}^6 \text{ sec}^{-1}. \quad (3.5)$$

The effect of the screening increases the reaction rate, but the correction is within the uncertainty in Γ_{E2} .

It is interesting to compare our result with Salpeter's.⁵⁾ At $T=2$ and $\rho=2.5 \times 10^4 \text{ g cm}^{-3}$ (3.5) and (2.18) gives us

$$\varepsilon = (\Gamma_{E2}/7 \times 10^{-10} \text{ Mev}) (\rho/2.5 \times 10^4)^2 1.5 \times 10^9 (T/2 \times 10^8 \text{ K})^{18.58} (x_{\text{He}})^3 \text{ erg g}^{-1} \text{ sec}^{-1}.$$

This is about 10^6 times larger than that given by Salpeter. Such a great difference is due mainly to the fact that he did not take the effect of the recently discovered resonance level in ^{12}C into account. This is one of the most important conclusions obtained in our work.



There are several levels of ^{16}O near the zero energy of the α -particle, as shown in Fig. 3. Among these levels the contributions from 7.12 Mev (1^-) and 6.91 Mev (2^+) levels need to be compared. The E1 radiation from the former level with zero isotopic spin would have a far greater width than the E2 and the E1-E3 cascade radiations from the latter, if the selection rule with regard to the isotopic spin were not taken into account. As the Z component of the isotopic spin is zero in the system concerned, the E1 transition is possible only because of the mixing¹¹⁾ of a state with isotopic spin unity. The probability for the mixing is believed as about one percent,¹¹⁾ so that the E1 width is smaller by about factor several hundred than in the case where such a selection rule is irrelevant. Even so the E1 width is nearly several times larger than the E2 width. Moreover, the resonance denominator in (2.13) is about four times larger for the 2^+ level than for the 1^- level, on account of $E_\eta = 0.2 \text{ Mev}$. Therefore, levels other than the 7.12 Mev 1^- level contribute to the reaction rate no larger than one percent. The E1 radiation width must be larger than 0.1 ev, because the lifetime of the level is observed as shorter than $8 \times 10^{-10} \text{ sec}$.¹¹⁾ Since the mixing of the isotopic spin states is regarded to be smaller than 1% it will be safe to limit the E1 width as

$$0.1 \text{ ev} < \Gamma_{E1} < 1 \text{ ev}. \quad (3.6)$$

For convenience we tentatively fix the value of

$$\Gamma_{E1} = 3.1 \times 10^{-7} \text{ Mev} \quad (3.6')$$

in the final formula.

Then the partial wave which contributes to the absorption of an α -particle is a P wave. The absorption width is calculated for $D \simeq 4 \text{ Mev}$ and $R = 5.4 \times 10^{-13} \text{ cm}$ as

$$\Gamma_{s,l=1}(E_\theta) = 1.62 \times 10^{+6} 10^{-20.12/T^{1/3}} \text{ Mev}, \quad (3.7)$$

where $E_g = 0.2$ Mev at 10^8 K. Hence the total width is practically equal to Γ_{E1} and is negligibly small compared with $E_g - E_n \simeq 0.278$ Mev at $T = 1.4 \times 10^8$ K. This allows us to use the non-resonance approximation (2.13) under the condition $E_g - E_n > \gamma_g \gtrsim kT \gg \Gamma'_n$.

Now we are able to give the reaction rate

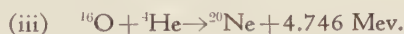
$$P(^{12}\text{C} + ^4\text{He} \rightarrow ^{16}\text{O}) = p_{c \rightarrow o} n_{\text{He}} n_c \quad (3.8)$$

with

$$p_{c \rightarrow o} = \frac{4.23 \times 10^{-13} 10^{-30.16/T^{1/8}}}{T^{2/3}} \left(\frac{\Gamma_{E1}}{3 \times 10^{-7} \text{ Mev}} \right) \text{ cm}^3 \text{ sec}^{-1}. \quad (3.9)$$

The reaction rate is much faster than that for $3 \cdot ^4\text{He} \rightarrow ^{12}\text{C}$ under the stellar conditions under consideration.

These formulas are valid for temperatures below 5.2×10^8 K, at which E_g lies at the midst between the 7.12 Mev level and the 8.6 Mev level. Above this temperature the contribution from the latter level has to be taken into account.



A recent experiment¹⁵⁾ has revealed a level scheme of ^{20}Ne different from that given in reference 11. The new level scheme is shown in Fig. 4. The most important level is the 4.95 Mev one, but its spin and parity are not yet undetermined. We calculated the radiation width, assuming the spin and parity of the level as 0^+ , 1^- or 2^+ . The result is shown in Table 2. Although its value strongly depends on the assumed spin and parity, the uncertainty does not come in the final formula, because of the following reasons.

Table 2. Radiation and electron emission widths (ev) for ^{20}Ne (4.95 Mev level)

| Γ_i | Assumed spin and parity of 4.95 Mev level | | |
|-----------------|---|--------------------------------|---|
| | 0^+ | 1^- | 2^+ |
| Γ_{E1} | 1.1 a (\rightarrow 4.2 Mev level, 1^-) | 61.4 a (\rightarrow ground) | 1.1 a (\rightarrow 4.2 Mev level, 1^- or 3^-) |
| Γ_{pair} | $\ll 10^{-6}$ (\rightarrow ground) | | 3.7×10^{-2} (\rightarrow 1.6 Mev level) |
| $\Gamma_{2\pi}$ | | | |
| Γ_K | | | |
| Γ_{E2} | 5.4×10^{-4} (\rightarrow 1.6 Mev level) | | |

The decay widths of possible modes are estimated for the 4.95 Mev level of ^{20}Ne , assuming its spin and parity as 0^+ , 1^- and 2^+ , and using Weisskopf's formula⁹⁾

The factor a in widths shows a constant. If the E1 transition is partially forbidden by the isospin selection rule, $a = 10^{-2} \sim 10^{-3}$, while otherwise $a = 1$.

At 10^8 K E_g is estimated as about 0.25 Mev and, consequently, $E_g - E_n = 0.05$ Mev. Since this is as small as γ_g , there appears a resonance peak in a broad Gamow peak. Hence we are allowed to use the resonance approximation, in which only a smaller one of radiation and absorption widths appears. The radiation width lies between 10^{-3} and

1 ev, as seen in Table 2, while the absorption width is estimated as smaller than 10^{-13} ev. The latter is expressed for the spin of the level, l , $R=5.9 \times 10^{-13}$ cm and $D \simeq 3$ Mev as

$$\Gamma_s(E_n) = 1.6 \times 10^{-19} \cdot 10^{-0.19l(l+1)} \text{ Mev.} \tag{3.10}$$

Now one sees that the absorption width is smaller than Γ' , and thus $\Gamma' \Gamma' / (\Gamma' + \Gamma') \simeq \Gamma'_s$ is a very good approximation.

Using the resonance approximation (2.14), the reaction rate is calculated as

$$P(^{16}\text{O} + ^4\text{He} \rightarrow ^{20}\text{Ne}) = p_{\text{O} \rightarrow \text{Ne}} n_{\text{He}} n_{\text{O}}, \tag{3.11}$$

with

$$p_{\text{O} \rightarrow \text{Ne}} = \frac{(2l+1)}{1.54^{l(l+1)}} \frac{2.19 \times 10^{-31} \cdot 10^{-10.98/T}}{T^{3/2}} \text{ cm}^3 \text{ sec}^{-1}. \tag{3.12}$$

This is rather insensitive to the value of l . The reaction rate is much larger than that for the formation of ^{16}O at $T < 2.6$. This is due to the recently observed level at 4.95 Mev.

Since the spin and parity of the 4.95 Mev level are unknown yet, this level might be ineffective to absorb an α -particle, though this is very unlikely in view of the nuclear structure of ^{20}Ne . In order to consider such a case as well as to see the competition with other levels, we calculate the reaction rates and the rates of energy generation due to 4.20 Mev and 5.62 Mev levels by using the non-resonance formula:

Table 3. Summary for the synthesis of ^{12}C , ^{16}O and ^{20}Ne

| Reaction | $3\ ^4\text{He} \rightarrow ^{12}\text{C} + \gamma$ | $^4\text{He} + ^{12}\text{C} \rightarrow ^{16}\text{O} + \gamma$ | $^4\text{He} + ^{16}\text{O} \rightarrow ^{20}\text{Ne} + \gamma$ |
|---|---|--|---|
| Q (Mev) | 7.278 | 7.149 | 4.746 |
| Main absorbing level | 7.65 Mev (0^+) | 7.12 Mev (1^-) | 4.95 Mev (?) |
| E_n (Mev) | 0.28 | -0.03 | 0.20 |
| E_g (Mev) | $0.146\ T^{2/3}$ | $0.199\ T^{2/3}$ | $0.247\ T^{2/3}$ |
| $ E_n - E_g $ at $T=1^*$ (Mev) | 0.13 | 0.23 | 0.05 |
| τ_g (Mev) | $0.0419\ T^{5/6}$ | $0.0489\ T^{5/6}$ | $0.0544\ T^{5/6}$ |
| Resonance or non-resonance | resonance at $0.7 \lesssim T \lesssim 30$ | non-resonance at $T \lesssim 5.2$ | resonance at $0.5 \lesssim T \lesssim 6$ |
| Γ_i (Mev) | $\Gamma_{E2} = (7 \pm 3) \times 10^{-10}$ | $\Gamma_{E1} = (0.1 \sim 1) \times 10^{-6}$ | $\Gamma_i = 10^{-9} \sim 10^{-6}$ |
| Γ_s (Mev) | $\Gamma_s(E_n) = 1.3 \times 10^{-6}$ | $\Gamma_s(E_g) = 1.6 \times 10^{+6} \times 10^{-20.12/T^{1/3}}$ | $\Gamma_s(E_n) = 1.6 \times 10^{-19} \times 10^{-0.19l(l+1)}$ |
| $P\ \text{sec}^{-1}\ g^{-1}$ | $\frac{3.66 \times 10^{18-18.75/T}}{T^3} \times \rho^2 x_H e^3$ | $\frac{3.16 \times 10^{33-30.15/T^{1/3}}}{T^{2/3}} \times \rho x_H e\ x_O$ | $\frac{(2l+1)}{1.54^{l(l+1)}} \frac{\rho x_H e\ x_O}{T^{3/2}} \times 1.23 \times 10^{15-10.98/T}$ |
| s (for $\varepsilon = a \times (T/1.4)^s$) | 27.8 | 20.0 | 15.4 |
| Uncertainty factor from Γ_i and Γ_s | ~ 2 | ~ 6 | ~ 2 |
| Contributions from other levels | 9.61 Mev level at $T \gtrsim 21$ | 8.6 Mev level at $T \gtrsim 5.2$ | 5.62 Mev level at $T \gtrsim 6$ |

*) T is measured in 10^8K . a is the value of ε at $T = 1.4 \times 10^8\text{K}$.

$$p_{O \rightarrow Ne} = \frac{(2l+1)}{1.54^{l(l+1)}} \left(\frac{I'_l}{10^{-6}} \right) \frac{1.65 \times 10^{-11} \cdot 10^{-37.85/T^{1/3}}}{T^{2/3}} \text{ cm}^3 \text{ sec}^{-1} \quad (3.13)$$

for the 4.20 Mev level;

$$p_{O \rightarrow Ne} = \frac{(2l+1)}{1.54^{l(l+1)}} \left(\frac{I'_l}{10^{-6}} \right) \frac{8.76 \times 10^{-12} \times 10^{-37.85/T^{1/3}}}{T^{2/3}} \text{ cm}^3 \text{ sec}^{-1} \quad (3.14)$$

for the 5.62 Mev level. In obtaining (3.13) and (3.14), we have evaluated the resonance denominator $(E_n - E_g)^{-2}$ in (2.15) with $E_0 \simeq E_g$ at $T = 1.4 \times 10^8$ K. l and I'_l

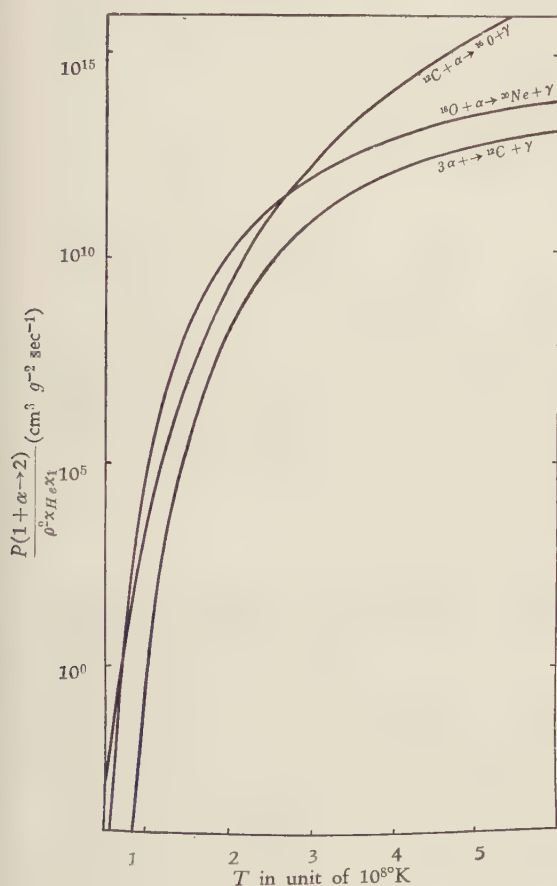


Fig. 5.

$$\frac{P(3\alpha \rightarrow {}^{12}\text{C})}{\rho^3 x_{He}^3} \times 10^3, \quad \frac{P({}^{12}\text{C} + \alpha \rightarrow {}^{16}\text{O})}{\rho^2 x_{He} x_{\alpha}} \\ \text{and} \quad \frac{P({}^{16}\text{O} + {}^4\text{He} \rightarrow {}^{20}\text{Ne})}{\rho^2 x_{He} x_{\alpha}}$$

The factor 10^3 in the case of $3 \cdot {}^4\text{He} \rightarrow {}^{12}\text{C}$ is applied for the central density of our star, 10^3 g cm^{-3} .

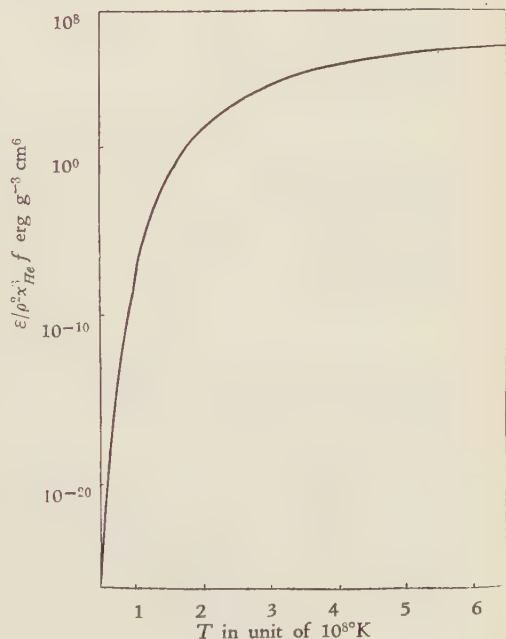


Fig. 6.

Energy generation per gram per second for the synthesis of ${}^{12}\text{C}$, ${}^{16}\text{O}$ and ${}^{20}\text{Ne}$.

This is obtained by multiplying 2.63 by the energy generation rate for the synthesis of ${}^{12}\text{C}$. Further correction will be made by multiplying a correction factor f shown in Fig. 10.

are unknown, but I' may be of the order of 1 to 10^{-3} ev in both cases. Even if we make maximum estimates of the reaction rates, the contributions from these levels are negligible in comparison with that from the 4.95 Mev level, as far as the latter level is effective for absorption; at temperatures above 6×10^8 K, the 5.62 Mev level becomes more important than the 4.95 Mev level.

Necessary data for the syntheses of ^{12}C , ^{16}O and ^{20}Ne are summarized in Table 3. In § 5 we shall see that the temperature and density of interest lie about $(1.3 \sim 1.5) \times 10^8$ K and $10^3 \sim 10^4$ g/cm³, respectively. Under such conditions, the formation rate of ^{20}Ne is much larger than those of ^{12}C and ^{16}O which are comparable, as seen in Fig. 5. Therefore, the whole series of the element synthesis is essentially determined by the formation rate of ^{12}C , most of ^{16}O formed from ^{12}C being quickly turned into ^{20}Ne . Hence the abundance of ^{20}Ne should be larger than that of ^{16}O , in contrast to the conclusion of Hoyle.⁶⁾ This problem will be studied in more detail in § 5.

The rate of energy generation due to the whole series is expressed as

$$\begin{aligned} \varepsilon &= (Q_{\alpha \rightarrow \text{C}} P_{\alpha \rightarrow \text{C}} + Q_{\text{C} \rightarrow \text{O}} P_{\text{C} \rightarrow \text{O}} + Q_{\text{O} \rightarrow \text{Ne}} P_{\text{O} \rightarrow \text{Ne}}) \rho^{-1} \\ &= f(Q_{\alpha \rightarrow \text{C}} + Q_{\text{C} \rightarrow \text{O}} + Q_{\text{O} \rightarrow \text{Ne}}) P_{\alpha \rightarrow \text{C}} \rho^{-1} = \frac{1.64 \times 10^{11-18.75/T}}{T^3} f x_{\text{He}}^3 \rho^2, \end{aligned} \quad (3.15)$$

$$f = \frac{Q_{\alpha \rightarrow \text{C}} x_{\text{He}}^2 + Q_{\text{C} \rightarrow \text{O}} p_{\text{C} \rightarrow \text{O}} x_{\text{C}} / p_{\alpha \rightarrow \text{C}} + Q_{\text{O} \rightarrow \text{Ne}} p_{\text{O} \rightarrow \text{Ne}} x_{\text{O}} / p_{\alpha \rightarrow \text{C}}}{(Q_{\alpha \rightarrow \text{C}} + Q_{\text{O} \rightarrow \text{Ne}} + Q_{\text{C} \rightarrow \text{O}}) x_{\text{He}}^2}. \quad (3.16)$$

Here f is a quantity of the order of unity which varies slowly as x_{He} , x_{O} and x_{Ne} change in the course of the element synthesis. The temperature dependence of $\varepsilon/f x_{\text{He}}^3 \rho^2$ is shown in Fig. 6. The exact value of this quantity will be calculated also in § 5 in connection with the problem of the element synthesis.

§ 4. Synthesis of ^{24}Mg to ^{40}Ca

^{24}Mg , ^{28}Si , ^{32}S , ^{36}Ar and ^{40}Ca can be synthesized by capturing α -particles successively from ^{20}Ne . The formation rates of these nuclei have never been observed. From experiments of neutron scattering, in which excited energies are higher by a few Mev than those of interest in the capture of α -particles, the average level distance is expected as about 0.1 Mev. At high temperatures this is smaller than the width of the Gamow peak, so that the statistical theory is applied to evaluate the absorption cross sections. In order to obtain the reaction rates applicable for a wide range of temperatures, however, more careful examination is required. (2.15) indicates that the ratio of the non-resonance to the resonance contributions is approximately given by

$$P_n/P_r \simeq \gamma_n I'/D^2. \quad (4.1)$$

Here the Gamow width γ_n is at most ten times the level distance $D \simeq 0.1$ Mev and the total width I' is practically equal to the radiation width that is at most 10^{-4} Mev. Hence (4.1) is estimated as not larger than 10^{-2} . Consequently, we have only to take account of the resonance contribution. As energy levels are unknown, the Gamow energy

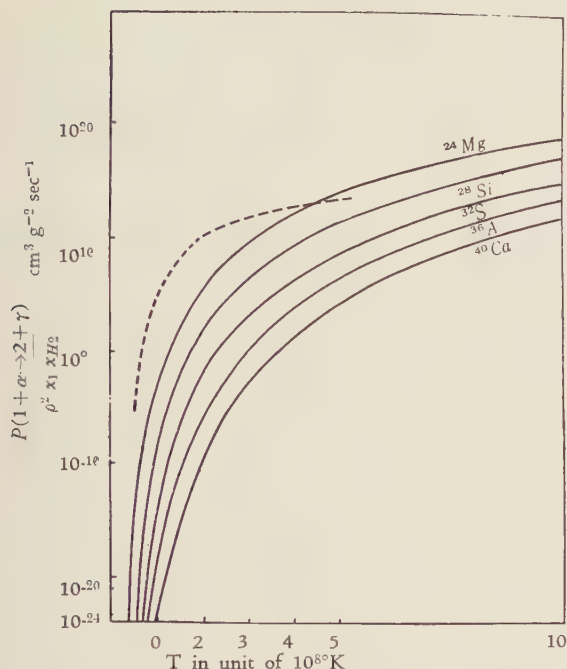


Fig. 7. Production rates of ^{24}Mg , ^{28}Si , ^{32}S , ^{36}A and ^{40}Ca $P(1 + ^4\text{He} \rightarrow 2 + \gamma) / \rho^2 x_1 x_{\text{He}}$.

The dotted curve shows $P(^{16}\text{O} + ^4\text{He} \rightarrow ^{20}\text{Ne}) / \rho^2 x_{\text{He}} x_{\text{Ne}}$ of ^{20}Ne with the assumed spin 0^+ of 4.95 Mev level.

may be substituted in place of E_n in the resonance formula.

$$P \simeq (2\pi / \mu k T)^{3/2} \hbar^2 n_1 n_2 \Gamma_s(E_\rho) \exp(-E_\rho / kT)$$

$$\simeq 8.0 \times 10^{-12} [(A_1 + A_2) / A_1 A_2]^{3/2}$$

$$n_1 n_2 I'_s(E_\rho) T^{-3/2} 10^{-5.1, 10 E_\rho / T}. \quad (4.2)$$

The formation rates of ^{24}Mg to ^{40}Ca are calculated by the use of (4.2) and are shown in Table 4 and Fig. 7. Comparing those rates with the formation rates of ^{12}C , ^{16}O and ^{20}Ne shown in Fig. 5, we find that the formation of ^{24}Mg and heavier elements in the synthesis of elements by helium capturing processes can be neglected in a temperature range $T \lesssim 4 \times 10^8 \text{K}$, unless the density is appreciably lower than 10^3g/cm^3 .

Table 4. Summary for the syntheses of ^{24}Mg to ^{40}Ca

| Reactions | $E_\rho(\text{Mev})$ | $r_\rho(\text{Mev})$ | Γ_s/D | $p_{1+\alpha \rightarrow 2} = \frac{P(1 + \text{He} \rightarrow 2)}{n_1 n_{\text{He}}}$ | $p_{1+\alpha \rightarrow 2} \simeq a \times (T/4)^s$ | |
|--|----------------------|----------------------|----------------------------|---|--|------|
| | | | | | a | s |
| $^{20}\text{Ne} + \alpha \rightarrow ^{24}\text{Mg}$ | $0.280 T^{2/3}$ | $0.059 T^{5/6}$ | $10^{8.30-29.80T^{-1/3}}$ | $\frac{2.6 \times 10^{-4-43.92T^{-1/3}}}{T^{3/2}}$ | 7.3×10^{-33} | 21.2 |
| $^{24}\text{Mg} + \alpha \rightarrow ^{28}\text{Si}$ | $0.331 T^{2/3}$ | $0.0629 T^{5/6}$ | $10^{9.42-33.41T^{-1/3}}$ | $\frac{3.6 \times 10^{-3-50.10T^{-1/3}}}{T^{3/2}}$ | 1.1×10^{-35} | 24.2 |
| $^{28}\text{Si} + \alpha \rightarrow ^{32}\text{S}$ | $0.369 T^{2/3}$ | $0.0665 T^{5/6}$ | $10^{10.49-37.55T^{-1/3}}$ | $\frac{4.0 \times 10^{-2-56.12T^{-1/3}}}{T^{3/2}}$ | 2.4×10^{-37} | 27.1 |
| $^{32}\text{S} + \alpha \rightarrow ^{36}\text{A}$ | $0.406 T^{2/3}$ | $0.0699 T^{5/6}$ | $10^{11.69-41.20T^{-1/3}}$ | $\frac{4.2 \times 10^{-1-61.66T^{-1/3}}}{T^{3/2}}$ | 8.1×10^{-41} | 30.1 |
| $^{36}\text{A} + \alpha \rightarrow ^{40}\text{Ca}$ | $0.441 T^{2/3}$ | $0.0726 T^{5/6}$ | $10^{12.47-44.70T^{-1/3}}$ | $\frac{4.0 \times 10^{-66.90T^{-1/3}}}{T^{3/2}}$ | 3.8×10^{-43} | 32.3 |

Production rates and necessary quantities for them are listed for ^{24}Mg , ^{28}Si , ^{32}S , ^{36}A and ^{40}Ca . T is expressed in unit of 10^8K .

§ 5. Application to astrophysical problems

The results obtained in the foregoing sections will be applied to estimate the temperature in the core of stars belonging to the horizontal branch in the H-R diagram and to calculate the relative abundances of ^{12}C , ^{16}O and ^{20}Ne which are formed in the interior of these stars.

A. The central temperature. In a recent study, Hoyle and Schwarzschild⁽¹⁶⁾ has shown that stars of globular clusters with mass slightly larger the solar mass evolve from the main sequence to the red giant within a time of about 6×10^8 years, increasing their central temperature. In the core of such stars hydrogen has been converted into helium. As the temperature in the core increases as high as 10^8 K, helium nuclei should begin to react with themselves, forming carbon. As soon as this reaction starts, the stars are presumed to evolve from right to left along the horizontal branch in the H-R diagram.^(16,17) The luminosity of stars in this branch is given by

$$L = (40 \sim 1600) L_{\odot} \quad (5.1)$$

corresponding to the absolute photo-visual magnitude $1 \sim -3$. This luminosity is to be explained in terms of the helium reactions which are assumed to be the main energy sources of these stars.

In order to derive the luminosity on the basis of the rate of energy generation obtained in § 3, one has to know the structure of the stars. Since the energy generation occurs only in a small central region on account of its steep temperature dependence, which is expressed from (3.15) as

$$\varepsilon = \varepsilon_0 x_{He}^3 f \rho^s T^s \quad (5.2)$$

with $s = 40 \sim 26$ for $T = (1.0 \sim 1.5) \times 10^8$, we can take a polytropic model with index n for the structure of this region. Then, neglecting the radiation pressure compared with the gas pressure and also the electron degeneracy, distributions of temperature and density are expressed in terms of the Emden function $\theta(\xi)$ as follows:⁽¹⁸⁾

$$T/T_c = \theta, \quad \rho/\rho_c = \theta^n, \quad \gamma = a\xi^2, \quad (5.3)$$

and

$$a^2 = (n+1)kT_c/4\pi G\mu m_H \rho_c,$$

where G is the constant of gravitation and μ the mean molecular weight of gas. For the present, T is expressed in ordinary unit. The total energy liberation is given by

$$L = \int \varepsilon 4\pi r^2 dr = 4\pi \varepsilon_0 x_{He}^3 f \rho_c^3 T_c^s a^3 \int \theta^{3n+s} \xi^2 d\xi. \quad (5.4)$$

Evaluating the last integral in (5.4) by approximating $\theta(\xi)$ for $\xi \ll 1$ as

$$\theta = 1 - \frac{1}{6} \xi^2 + \frac{n}{120} \xi^4 - \dots = e^{-\xi^2/6} \left\{ 1 + \left(\frac{n}{120} - \frac{1}{72} \right) \xi^4 + \dots \right\} \quad (5.5)$$

and by extending the integration limit to $\xi = \infty$, which produce no appreciable errors for $3n+s \gg 1$, we find

$$L \simeq \varepsilon_0 x_{He}^3 f \left(\frac{3n+3}{3n+s} \frac{k}{2G\mu m_H} \right)^{3/2} \rho_c^{3/2} T_c^{s+3/2}. \quad (5.6)$$

Noting that (5.6) is rather insensitive to the value of n so long as $s \gg n$, we take as the appropriate values of n and s

$$n=1.5 \text{ and } s=28, \quad (5.7)$$

which correspond to the polytropic index for the convective equilibrium and to the temperature 1.4×10^8 K, respectively. Further, expressing T_c in unit of 10^8 K and putting

$$\rho_c = 10^3 \alpha T_c^3, \quad (5.8)$$

we have from (5.6)

$$L/L_\odot \simeq 290 (\alpha/\mu)^{3/2} x_{He}^3 f (T_c/1.4)^{34}. \quad (5.9)$$

The value of α defined by (5.8) depends on the whole structure of stars, but it is known to be of the order of unity for stellar model of helium burning stars studied preliminarily by Hoyle and Schwarzschild.¹⁶⁾ The determination of T_c from (5.9) is not sensitive to the value of $(\alpha/\mu)^{3/2} x_{He}^3 f$ so that we fix it as

$$(\alpha/\mu)^{3/2} x_{He}^3 f = 1 \quad (5.10)$$

on account of $\alpha \gtrsim 1$, $x_{He} \lesssim 1$, $f \simeq 1$ and $2 > \mu \gtrsim 4/3$. Then, we find from (5.1) and (5.9)

$$T_c = (1.40 \pm 0.10) \times 10^8 \text{ K} \quad (5.11)$$

for the stars belonging to the horizontal branch. Since the rate of energy generation around this temperature depends very steeply on T_c , T_c is determined within a rather narrow range, even if other quantities α and x_{He} are varied in a wide range. For example, decreases of $(\alpha x_{He}^2/\mu)^{3/2} f$ by factor 10 and 10^3 results in the increases of T_c by 7 percent and 20 percent, respectively.

As has been stated in § 3, the value of T_c obtained above is far smaller than that by Salpeter.⁵⁾ This is very important in discussing the synthesis of elements in the stars, as shown in what follows.

B. *Abundances of ^4He , ^{12}C , ^{16}O and ^{20}Ne .* Since the central temperature is determined as rather low, the formation of ^{24}Mg and heavier nuclei can be neglected in discussing the synthesis of elements in the stars under consideration.

The time variation of relative abundances is expressed by the following set of equations.

$$dx_{He}/dt = -3 (\rho/M_\alpha)^2 p_{3\alpha \rightarrow C} x_{He} (x_{He}^2 + \kappa x_C/3 + \kappa' x_O/4), \quad (5.12a)$$

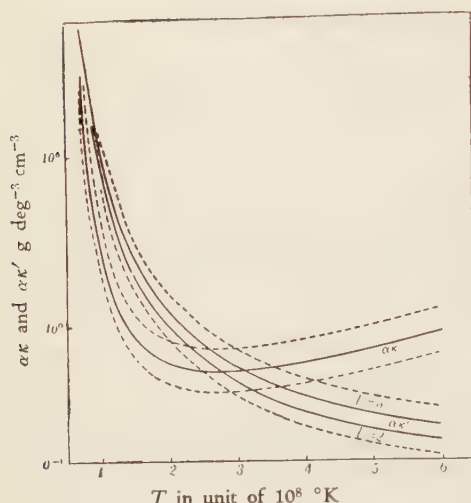
$$dx_C/dt = 3 (\rho/M_\alpha)^2 p_{3\alpha \rightarrow C} x_{He} (x_{He}^2 - \kappa x_C), \quad (5.12b)$$

$$dx_O/dt = 3 (\rho/M_\alpha)^2 p_{3\alpha \rightarrow C} x_{He} ((4/3) \kappa x_C - \kappa' x_O), \quad (5.12c)$$

$$\text{and} \quad dx_{Ne}/dt = 3 (\rho/M_\alpha)^2 p_{3\alpha \rightarrow C} x_{He} (5/4) \kappa' x_O. \quad (5.12d)$$

Here κ and κ' are dimensionless quantities defined by

$$\kappa = \frac{p_{C \rightarrow O} M_\alpha}{3 p_{3\alpha \rightarrow C} \rho} \quad \text{and} \quad \kappa' = \frac{p_{O \rightarrow Ne} M_\alpha}{3 p_{3\alpha \rightarrow C} \rho}. \quad (5.13)$$

Fig. 8. $\alpha\kappa$ and $\alpha\kappa'$

Dashed curves show their domain of uncertainties arising from Γ_3 and Γ_8 . l indicates the assumed spin of 4.95 MeV level of ^{20}Ne which has not been determined yet.

unity and (5.12) are solved as

$$x_O \simeq \nu t, \quad x_O \simeq \kappa (\nu t)^2 \quad \text{and} \quad x_{Ne} \simeq \kappa \kappa' (\nu t)^3. \quad (5.15)$$

At $t \simeq 1/\nu\kappa'$, (5.15) shows that we have $\kappa x_O \simeq \kappa' x_O$ and $x_O \simeq x_{Ne}$ in order of magnitude. For $t \gtrsim 1/\nu\kappa'$ we can put an approximate relation,

$$4/3 \cdot \kappa x_O \simeq \kappa' x_O, \quad (5.16)$$

on account of the high values of κ' and κ'/κ . The relation (5.16) states that the number of ^{16}O formed from ^{12}C per unit time is nearly equal to that converted into ^{20}Ne . Solving again (5.12) with (5.16) for $1/\nu\kappa' \lesssim t \lesssim 1/\nu$, we have in order of magnitude

$$x_O \simeq \nu t, \quad x_O \simeq (\kappa/\kappa') \nu t \quad \text{and} \quad x_{Ne} \simeq \kappa (\nu t)^2. \quad (5.17)$$

It is found in this way that the abundance of ^{16}O remains very small compared with that of ^{12}C and also that of ^{20}Ne except for a short period $t \lesssim 1/\nu\kappa'$. This result is in contradiction with that of Hoyle⁽¹⁾ who discarded the formation of ^{20}Ne and determined the stellar temperature based on the cosmic abundance of ^{12}C and ^{16}O .

A more detailed study of the relative abundances is carried out by the numerical solution of (5.12). For this purpose we eliminate time t from (5.12) and obtain a set of equations for x_O , x_O and x_{Ne} against x_{He} .

$$dx_O/dx_{He} = - \frac{x_{He}^2 - \kappa x_O}{x_{He}^2 + (\kappa/3)x_O + (\kappa'/4)x_O}, \quad (5.18a)$$

$$dx_{Ne}/dx_{He} = - \frac{(4/3)\kappa x_O - \kappa' x_O}{x_{He}^2 + (\kappa/3)x_O + (\kappa'/4)x_O}, \quad (5.18b)$$

The values of $\alpha\kappa$ and $\alpha\kappa'$ are represented in Fig. 8 with possible uncertainties. As the temperature specifying κ and κ' , we can take the central one on account of the steep temperature dependence of reaction rates. Indeed, the integrand in (5.4) has its maximum at $1-\theta \equiv 1-T/T_0 = 1/(3n+s) \simeq 1/32.5$. Then, we have $\kappa \simeq 1$ and $\kappa' \simeq 100$, if we adopt the most probable values represented by solid curves in Fig. 8. At the same temperature (3.5) leads us to

$$\nu \equiv 3 (\rho/M_a) p_{3\alpha \rightarrow O} \simeq 10^{-11} \text{ sec}^{-1}. \quad (5.14)$$

We shall first examine the qualitative nature of the solutions of (5.12), assuming that changes of κ and κ' are small in the course of the element synthesis. For $t < 1/\nu\kappa' \sim 10^{12} \text{ sec}$, x_{He} can be regarded as

and

$$dx_{Ne}/dx_{He} = - \frac{(5/4) \kappa' x_O}{x_{He}^2 + (\kappa/3) x_C + (\kappa'/4) x_O} \tag{5.18c}$$

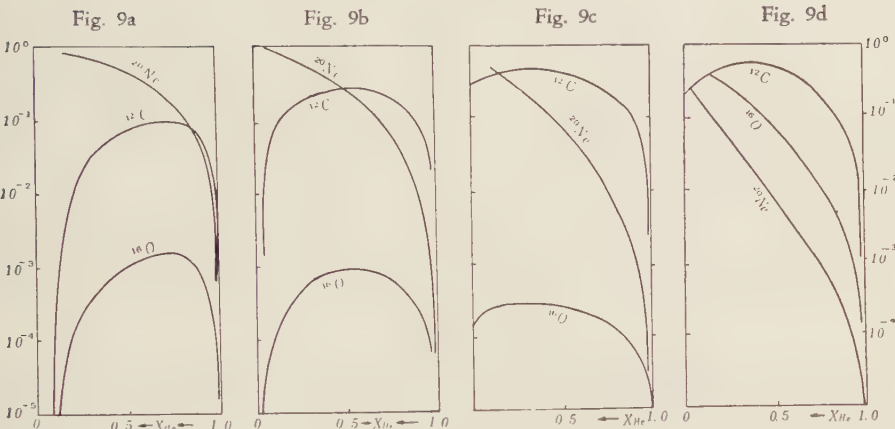


Fig. 9. Relative abundances of ^{12}C , ^{16}O and ^{20}Ne κ and κ' are chosen respectively as follows:

| | κ | κ' |
|---|----------|-----------|
| a | 5 | 250 |
| b | 1 | 250 |
| c | 0.2 | 250 |
| d | 0.2 | 0.1 |

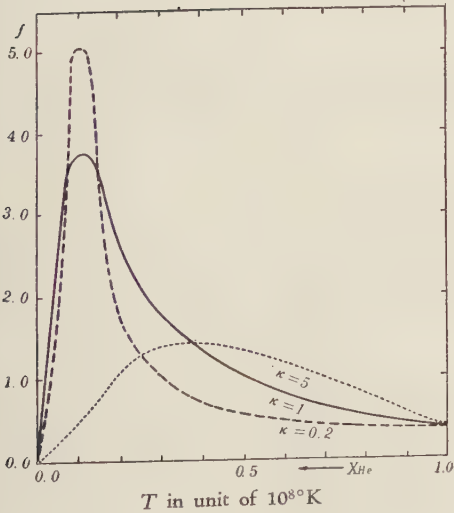


Fig. 10. Correction factor for the energy generation

very small, say, ten percent. Therefore, results obtained above are not altered essentially if the variations of κ and κ' are taken into account in solving (5.18).

The solutions of these equations are given in Fig. 9 for a number of given sets of κ and κ' . It is found in the interesting cases, $\kappa' \gg 1$ and $\kappa' \gg \kappa$, that the final abundance of ^{16}O is very small compared with those of ^{12}C and ^{20}Ne and that the final ratio x_C/x_{Ne} is less than 0.6 percent for $\kappa \gtrsim 0.6$. In the same interesting case, the factor (3.16) in the total energy generation is also plotted against x_{He} in Fig. 10.

As the helium reactions proceed with x_{He} decreasing, the central temperature of the stars must increase in order to give the nearly constant luminosity of the horizontal branch. However, (5.9) together with Fig. 10 shows that the required increase in temperature is rather small unless x_{He} becomes

At $T \approx 1.4 \times 10^8 K$, most of helium nuclei are found to be converted into ^{20}Ne , in definite disagreement with the observed cosmic abundances. It will not be correct, however, if one thinks that the cosmic abundances of ^{12}C , ^{16}O and ^{20}Ne are accounted for merely in terms of the helium capturing processes. For the elements thus formed will be converted into heavier ones in the later stages of stellar evolution. For example, $^{20}\text{Ne}(\alpha, \gamma)^{24}\text{Mg}$ and $^{20}\text{Ne}(p, \gamma)^{21}\text{Na}$ will be important processes to eliminate abundant ^{20}Ne , if there occurs some mixing of the chemical composition. $^{21}\text{Na}(\beta^+ \nu)^{21}\text{Ne}$ and $^{21}\text{Ne}(\alpha, n)^{24}\text{Mg}$ are considered as sources of building up heavier elements by W.A. Fowler et al.¹⁰⁾ On the other hand, if there is no mixing $^{20}\text{Ne} + ^{20}\text{Ne} \rightarrow ^{40}\text{Ca} + \gamma$ will be expected at the temperature $\gtrsim 10^9 K$. The heavy elements thus built up may be broken up through spallation reactions caused by energetic particles, when stellar matter is scattered into the interstellar space.²⁰⁾²¹⁾

§ 6. Conclusions

We have calculated the rates of thermonuclear reactions caused by successive helium capturing processes in the helium core. The rates obtained by us are much faster than those obtained previously by Salpeter⁵⁾ and Hoyle.⁶⁾ This is mainly because the resonance levels are discovered for ^{12}C and ^{20}Ne . Our results are in essential agreement with those calculated by Obi et al.⁸⁾

If the rate of energy generation is applied to interpret the luminosity of the stars belonging to the horizontal giant branch in the H-R diagram, the central temperature is estimated as $(1.4 \pm 0.10) \times 10^8 K$ for the density of about 10^3 g cm^{-3} . At about this temperature the formation of ^{24}Mg and heavier nuclei can be neglected and there occurs a series of syntheses from helium to neon. In this series the formation of ^{12}C from three ^4He nuclei is the slowest reaction, so that this determines the reaction rate of this series. The lifetime for the giant stars to reach the stage of R.R. Lyrae type variable stars is estimated as about 10^7 years. In this course of evolution ^{20}Ne is formed more abundant than others. This suggests that the element abundances determined by the helium capturing processes are considerably modified in later stages of evolution.

Our result can also be applied for the primeval synthesis of elements in the α - β - γ theory. The most serious difficulty in their theory lies in the point that the successive formation of heavier elements is locked at unstable nuclei with mass numbers five and eight. However, the large reaction rate of forming ^{12}C from three α -particles may remove such a difficulty. This problem will be discussed in a separate paper by one of us (C. H.) and his collaborators.

We should like to express our sincerest gratitude to Professors T. Hatanaka, Z. Hitotuyanagi and M. Taketani; their view of the stellar evolution motivated us to work in this subject. We are also grateful to Professor S. Nakamura and his collaborators, with whom we have had constant discussions on the same problem.

References

- 1) H. Bethe, Phys. Rev. **55** (1939), 434; Ap. J. **92** (1940), 118.
- 2) H. Bethe and C. R. Critchfield, Phys. Rev. **54** (1938), 248.
E. E. Salpeter, Phys. Rev. **88** (1952), 547.
- 3) W. A. Fowler, Proceedings of the Symposium on Nuclear Processes in Celestial Objects, Liege (1954), 88.
E. E. Salpeter, Phys. Rev. **97** (1955), 1237.
- 4) M. Taketani, T. Hatanaka and S. Obi, Prog. Theor. Phys. **15** (1956), 89; earlier papers are cited there.
- 5) E. E. Salpeter, Ap. J. **115** (1952), 326.
E. E. Salpeter, Ann. Rev. Nucl. Sci. **2** (1953), 4.
- 6) F. Hoyle, Ap. J. Supple. **1**, No. 2 (1954), 121.
- 7) E. E. Salpeter, Phys. Rev. **98** (1955), 1183. The authors are indebted to Professor Salpeter who let them know his calculation of the radiation widths concerning the helium capturing processes.
- 8) The same attempt was carried out by S. Obi and his collaborators. Prog. Theor. Phys. **16** (1956), 389.
With these authors we have made a close contact and enjoyed stimulating discussions.
- 9) J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics* (1952).
- 10) E. E. Salpeter, Aust. J. Phys. **7** (1954), 373.
- 11) F. Ajzenberg and T. Lauritzen, Rev. Mod. Phys. **27** (1955), 77.
- 12) V. K. Rasmussen, D. W. Miller and M. B. Sampson, Phys. Rev. **100** (1955), 181.
J. Seed, Phil. Mag. **46** (1955), 100. Harries, Proc. Phys. Soc. **67A** (1954), 153.
- 13) L. I. Schiff, Phys. Rev. **98** (1955), 1281.
- 14) D. G. Ravenhall, private communication. Devons, Manning and Towle, Proc. Phys. Soc. **69A**, No. 434A, 173. Others cited in ref. 11).
- 15) R. G. Freemantle, D. F. Prouse, A. Hossian and J. Rotblat, Phys. Rev. **96** (1955), 1270.
G. Shrank and G. K. O'Neill, Bull. Am. Phys. Soc. **1**(1956), No. 1, 29.
- 16) F. Hoyle and M. Schwarzschild, Ap. J. Supple. **2**, No. 2 (1955).
- 17) T. Hatanaka and S. Obi, private communication.
- 18) See, for example: S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Chicago (1939).
- 19) W. A. Fowler, G. R. Burbidge and E. M. Burbidge, Ap. J. **122** (1955), 271.
- 20) W. A. Fowler, G. R. Burbidge and E. M. Burbidge, Ap. J. Suppl. No. **17** (1955) 167.
- 21) S. Hayakawa, Prog. Theor. Phys. **15** (1956), 111.

Letters to the Editor

Charge Independence Hypothesis and Leptons

Tadashi Ouchi and Kei Senba

*Department of Physics, Hiroshima University,
Hiroshima*

March 9, 1956

The purpose of this note is to present a tentative scheme for leptons on the basis of the Pais-Nishijima-Gell-Mann theory¹⁾ to find out some good selection rules concerning lepton processes. One of the most important features of the P-N-G theory is in our opinion the assignment of η -charge to each hyperon and meson co-operated with the conservation laws between the charge independent strong interactions. If η -charge really has the sense of "attribute" for the leptons as Sachs²⁾'s anticipation, we are obliged to look for the charge independent interactions hidden among the lepton processes to assign η -charge to leptons. Fortunately Cini and Gamba³⁾ have found the interactions between nucleon and leptons apparently charge independent, if two kinds of neutrinos such as μ_0 and e_0 are introduced. Moreover many authors⁴⁾⁵⁾ have pointed out that there exists some regularity among the weak interactions realized in nature. Thus we may postulate, taking into account the study of D'Espagnat and Prentki,⁶⁾ that all of the fundamental particles have the definite transformation properties in the isotopic spin space and that all the elementary interactions concerning leptons are

charge independent. And we suppose that all deviations from the charge independence are due to electromagnetic effects and therefore slight. According to the research made by several authors assuming the interaction form as $e\bar{\psi}_a\gamma_\mu\psi_aA_\mu$ for all the fermions, this seems to be a reasonable possibility. But it remains unexamined whether the consistent interpretation of the mass difference between e and e_0 would be given or not.

Introducing two kinds of neutrinos (e_0 and ν) with vanishingly small rest masses, we assign to each particle the definite transformation property in Lorentz- and isotopic spin space and η -charge as listed in Table.

Table I. Types of (normal) fields

| Name | Lorentz space | Isotopic spin space | η -charge |
|---------|---------------|---------------------|----------------|
| μ^* | spinor | pseudo-vector | -1 |
| e | spinor | spinor of 1st kind | 0 |
| ν | spinor | scalar | -1 |
| K | boson | spinor of 1st kind | 1 |
| B_1 | boson | spinor of 1st kind | 1 |
| B_2 | boson | pseudo-vector | 0 |

* The existence of metastable μ_0 meson does not cause any trouble.

If the weak lepton interactions are assumed to be deduced by the true scalar Lagrangian having only the Yukawa-type interactions $(\bar{H}H)B + (\bar{L}L)B$, where H , L and B denote hyperons, leptons and the auxiliary boson fields, which are required to guarantee the charge conservation of the transition processes automatically⁶⁾, we can derive the conservation laws such as conservation of iso-fermions and fermions

(leptons and hyperons) by the same procedure as D'Espagnat and Prentki used. These B fields should be considered, if they really exist, to be of the similar properties as those proposed by Ogawa,⁵⁾ which have the masses heavier than about $1000 m_e$ and the interaction constants of the intermediate magnitude for all fermions in accordance with the observed universality of weak interactions and thus are hardly discovered. In applying our scheme to classification of the existing lepton processes, it is necessary to distinguish e_0 from ν in each process in such a way as ;

$$\begin{aligned} N \rightarrow P + e^- + e_0 & \quad \pi^- \rightarrow \mu^- + \nu \\ \mu^- + P \rightarrow N + \nu & \quad K^+ \rightarrow \mu^+ + e_0 + \pi^0. \end{aligned} \quad (\text{A})$$

Then by the use of the conservation laws we can deduce the following selection rules for the lepton processes ; (1) transitions having $\Delta\eta=0$ are allowed and (2) transitions having $\Delta\eta=1$ should be forbidden. Thus our scheme with two kinds of neutrinos produces the good selection rules except that the following processes cannot be ruled out,

$$\begin{aligned} K^+ \rightarrow e^+ + \nu, & \quad K^+ \rightarrow \pi^+ + e^+ + \mu^-, \\ \pi^+ \rightarrow e^+ + e_0, & \quad K^+ \rightarrow e^+ + \mu_0, \end{aligned} \quad (\text{B})$$

and several μ - and β -decays of hyperons.

Finally let us show that processes (B) do not present any difficulty to charge independence hypothesis for leptons by taking tentatively as the primary interactions the universal Fermi ones, which are limited by our Lagrangian and can change their coupling form from one process to another, keeping their coupling constants in the same order of magnitude. As was often remarked,⁴⁾⁵⁾ π -meson can decay not into e and e_0 but into μ and ν if β -decay

of nucleon contains only S- and T-couplings and $(n, p) - (\mu, \nu)$ interaction has PV coupling etc. This example suggests that the similar prediction is possible for the decay processes of K -meson if the coupling forms of μ - and β -decays of hyperons are suitably chosen. As our decay scheme for K -meson corresponds to the case b) considered by Costa and Dallaporta,⁷⁾ it will be evident that the consistent interpretation of $K_{\pi 2}$, $K_{\pi 3}$, $K_{\mu 3}$ and $K_{e 3}$ -decays is possible and that μ - and β -decays of hyperons have too long lifetimes compared with the normal decays of hyperons to be observed. The unwanted processes $K^+ \rightarrow \pi^+ + e^+ + \mu^-$ and $K^+ \rightarrow e^+ + \mu_0$ can be forbidden if we assign the types in space reflection⁸⁾ to each fermion as follows ;

$$\Xi N \in A, \Sigma \Lambda \in B, e^+ e_0 \mu^+ \mu_0 \in C, \mu^- \nu \in D, \quad (\text{C})$$

where ϵ reads "belong to type." The parity of μ^- is taken as conjugate of μ^+ to distinguish μ^- from the antiparticle $\mu^-(\epsilon C)$ of μ^+ and such a consideration is not required, but automatically satisfied, to $\Sigma(\epsilon B)^*$, which exhibits a striking contrast between C or D type fields and A or B type one. On the other hand the consistent explanation of the observed decay process can be expected due to the existence of the so-called universal Fermi interaction. Thus the selection rules obtained can be expressed briefly after Sachs in that the weak lepton interactions are limited to involve at least one neutrino ν or e_0 . The arguments so far given will be enough to show that the consistent lepton scheme with P-N-G theory is in

* This may be regarded as an extended use of the concept of parity doublets first introduced by Lee and Yang (Phys. Rev. 102 (1956), 290) and as consistent with our assumption.

is in principle possible* and compatible with the existing experiments. Further discussions including the electromagnetic interactions** will be given in detail in future elsewhere.

In conclusion the authors wish to express their sincere thanks to Prof. K. Sakuma, Dr. S. Ogawa and Mr. M. Yonezawa for their useful discussions.

- 1) A. Pais, *Physica* **19** (1953), 869; *Proc. Nat. Acad. Sci. U.S.* **40** (1954), 484; M. Gell-Mann, *Phys. Rev.* **92** (1953), 833; K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285.
- 2) R. G. Sachs, *Phys. Rev.* **99** (1955), 1573.
- 3) M. Cini and A. Gamba, *Nuovo Cimento* **10** (1953), 1040.
- 4) K. Iwata, S. Ogawa, H. Okonogi, B. Sakita and S. Oneda, *Prog. Theor. Phys.* **13** (1955), 19.
- 5) S. Ogawa, *Prog. Theor. Phys.* **15** (1956), 487.
- 6) B. D'Espagnat and J. Prentki, *Nuclear Physics* **1** (1956), 33.
- 7) G. Costa and N. Dallaporta, *Nuovo Cimento* **2** (1955), 519; N. Dallaporta, *ibid.* **1** (1955), 962. See also the paper cited in 4.
- 8) C. N. Yang and J. Tiomno, *Phys. Rev.* **79** (1950), 495.

* S. Gotō has shown in his private communication that the unwanted processes such as (B) can be diminished with the same selection rules by taking another scheme which is made of neutrino ν and Koropinski-Mahmoud scheme (*Phys. Rev.* **92** (1953), 1045).

** After this letter was submitted to *Progress of Theoretical Physics*, a paper by F. Duimio (*Nuovo Cimento* **2** (1955), 1308) came to our notice where the detailed discussions on K -meson decays are given, which seem to agree with our results.

Integral over Path Formulation of Statistical Theory of Irreversible Processes

David Falkoff

Brandeis University, Waltham,
Massachusetts, U.S.A.

September 4, 1956

A statistical theory of irreversible processes has been proposed by Onsager and Machlup¹⁾ and, independently, by Hashitsume.^{2), 3), 4)} The mathematical structure is similar to that used in the theory of Brownian motion^{5), 6)} and of thermal noise in linear passive electrical networks⁷⁾ in that one deals with linear stochastic differential equations, and the resulting probability distributions are stationary, Gaussian and Markoff.⁷⁾ Thus, for the one-dimensional case, the conditional probability for a fluctuating variable $\alpha(t)$ to have the value α'' at t'' , given that it had the value α' at $t' = t'' - \tau$, $\tau > 0$, has the form*

$$P_{OU}(\alpha' t' | \alpha'' t'') = \left[\frac{\Lambda}{\pi \sigma^2 (1 - e^{-2\Lambda\tau})} \right]^{1/2} \exp \frac{-\Lambda(\alpha'' - \alpha' e^{-\Lambda\tau})^2}{\sigma^2 (1 - e^{-2\Lambda\tau})}. \quad (1)$$

In the application to irreversible processes, $\Lambda = s/R$, $\sigma^2 = 2k/R$ where k is Boltzmann's constant and R and s are parameters relating to dissipation and entropy in the phenomenological equations. Onsager and Machlup have shown that (1) can also be cast in the elegant variational form:

* The subscript OU refers to Ornstein and Uhlenbeck (ref. 5) who first obtained this solution in their extension of Einstein's theory of Brownian motion (ref. 6).

$$P_{ov} = N(\tau) \exp \frac{-1}{2k} \left\{ \int_{\nu'}^{\nu''} \mathcal{L}^s(\alpha, \dot{\alpha}) dt \right\}_{\min} \quad (2)$$

where $N(\tau)$ is the same normalizing coefficient of the exponential as in (1) and $\mathcal{L}^s(\alpha, \dot{\alpha})$ is the "thermodynamic Lagrangian"

$$\mathcal{L}^s(\alpha, \dot{\alpha}) = \frac{R}{2} \left(\dot{\alpha} + \frac{s}{R} \alpha \right)^2 \quad (3)$$

They did not attempt to paths $\alpha(t)$ for continuous range of t since the measure of a single such path is zero. However, Hashitsume⁴⁾ (see also Falkoff, ref. 8) has shown that by making the heuristic assumptions a) that one can assign a probability to each path, $\alpha(t)$,

$$\text{prob}\{\text{path}, \alpha(t)\} \sim \exp \frac{-1}{2K} \int_{\nu'}^{\nu''} \mathcal{L}^s(\alpha, \dot{\alpha}) dt \quad (4)$$

and b) that the transition probability from α' at t' to α'' at t'' can be gotten by summing over all paths having these initial and final values, one may again obtain (1). His proof uses the Fourier series "method of Rice"⁹⁾ due to Einstein¹⁰⁾ which assumes the process is Gaussian.*

In the present note we shall explicitly relate the integral over path formulation of irreversible processes to Wiener integrals.¹¹⁾ In so doing the following insights are gained:

- i) the Wiener integral provides a rigorous mathematical means for expressing integral over fluctuation paths; at the same time the pathological nature of these paths is put quantitatively in evidence.
- ii) the Wiener integral may also be applied to obtain expectations of

functionals of paths.

- iii) it provides a rigorous classical analogue of Feynman's integral over path form of quantum mechanics.¹²⁾
- iv) it provides the means to extend the statistical theory of irreversible processes to non-Gaussian but Markoff processes corresponding to non-linear phenomenological equations.
- v) the perturbation expansions of quantum mechanics may be adapted to irreversible processes with Boltzmann's constant k , playing the role of $\hbar/2$.

To make more explicit points i and iii we state the Theorem:

$$P_{ov} = \int \mathcal{D}(\text{paths}) \text{prob}(\text{path}) \\ = \left[\exp \frac{-A}{\sigma^2} \int_{\alpha'}^{\alpha''} \alpha d\alpha \right] \left[\int_{\alpha'|\alpha''}^{\nu''} d_{\nu'} \alpha \right. \\ \left. \times \left\{ \exp \frac{-A^2}{2\sigma^2} \int_{\nu'}^{\nu''} \alpha^2 dt \right\} \right]. \quad (5)$$

The intermediate expression is to be interpreted in the same way as Feynman's integral over paths, the probability of a path being given by (4). In the last mathematically rigorous form the second square bracket is a generalization of a much treated Wiener integral^{13), 14)} in that here both initial and final values of α are fixed.** It has the value:¹⁴⁾

$$\left(\frac{A}{2\pi\sigma^2 \sinh A\tau} \right)^{1/2} \\ \cdot \exp \frac{-A[(\alpha'^2 + \alpha''^2) \cosh A\tau - 2\alpha'\alpha'']}{2\sigma^2 \sinh A\tau}. \quad (6)$$

This is exactly the same as the transformation amplitude $(\alpha't'|\alpha''t'')$, for the

* A similar proof has also been given in unpublished work by L. Tisza and I. Manning.

** (6) reduces to the Wiener integral treated in ref. 13 when $\alpha' = 0$, $\sigma^2 = 1/2$ and the integration over α'' is performed.

quantum mechanical simple harmonic oscillator^{12),15),16)} if one sets $\alpha = x$, $A = i\omega$ and $\sigma^2 = i\hbar/m$. Thus the formal correspondence between parameters of the irreversible process and the harmonic oscillator is: $2k \longleftrightarrow i\hbar$, $R \longleftrightarrow m$, $s/R \longleftrightarrow i\omega$. This connection is not unexpected¹⁴⁾ in virtue of the "Markoff structure" for the transition amplitudes in quantum mechanics and the similarity of the harmonic oscillator Lagrangian to (3). Indeed one may also evaluate more general Wiener integrals corresponding to Lagrangians other than (3) by the techniques of quantum mechanics, e.g. as Green's functions for partial differential equations analogous to the Schrödinger equation.

More surprising is the fact that in the first bracket in (5) one must use the relation

$$\int_{\alpha'}^{\alpha''} \alpha d\alpha = \frac{\alpha'^{1/2} - \alpha''^{1/2}}{2} - \frac{\sigma^2 \tau}{2}. \quad (7)$$

The extra term, $\sigma^2 \tau / 2$, appears because the random functions, $\alpha(t)$, are not of bounded variation and hence the ordinary integrals do not exist. However, a "stochastic integral" may be uniquely defined¹⁷⁾ for such functions and yields the unexpected result (7). Using (7) and (6) in (5) one obtains (1).

A more detailed account of the above points, with applications, is now in preparation.

- 1) L. Onsager and S. Machlup, *Phys. Rev.* **91** (1953), 1505.
- 2) N. Hashitsume, *Prog. Theor. Phys.* **8** (1952), 461.
- 3) N. Hashitsume, *Proc. Intern. Conf. Theor. Physics, Kyoto, 1953*, p. 495.
- 4) N. Hashitsume, *Prog. Theor. Phys.* **15** (1956), 369.
- 5) L. S. Ornstein and G. E. Uhlenbeck, *Phys. Rev.* **36** (1930), 823.

- 6) A. Einstein, *Ann. d. Physik* **17** (1905), 549.
- 7) M. C. Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.* **17** (1945), 323.
- 8) D. L. Falkoff, *Bull. Am. Phys. Soc.* **11** **1** (1956), 65.
- 9) S. O. Rice, *Bell Sys. Tech. Jour.* **23** (1944), 1.
- 10) A. Einstein, *Ann. d. Physik* **47** (1915), 879.
- 11) N. Wiener, *Acta Mathematica* **55** (1930), 117.
- 12) R. P. Feynman, *Rev. Mod. Phys.* **20** (1948), 367.
- 13) R. H. Cameron and W. T. Martin, *Bull. Am. Math. Soc.* **51** (1945), 73.
- 14) E. Montroll, *Comm. Pure Appl. Math.* **5** (1952), 415.
- 15) M. S. Partlett and J. E. Moyal, *Proc. Camb. Phil. Soc.* **45** (1948), 545.
- 16) K. Husimi, *Prog. Theor. Phys.* **9** (1953), 381.
- 17) J. L. Doob, *Stochastic Processes* (J. Wiley and Sons, New York) 1953 p. 443.

Note inserted in proof: Similar points of view are developed in an article by N. Saito and M. Namiki, *Prog. Theor. Phys.* **16** (1956), 71,

On the Bethe-Salpeter Equation for the Scalar-Spinor Particle System

Reiji Sugano and Yasuo Munakata

*Department of Physics, Kyoto University,
Kyoto*

August 25, 1956

The relativistic two-body problem can be treated field-theoretically in terms of the Bethe-Salpeter (B-S) equation.¹⁾ A recent study made by Wick²⁾ shed (some) light on the properties of this equation, especially on the boundary condition at $t = \pm \infty$ (t being the "relative" time variable). He showed that B-S equation in Lorentz space can be transformed into that in Euclidian space and solved the equation for the scalar-scalar system in the case of "ladder approximation" thus getting the discrete eigenvalues.³⁾ On the other hand,

the analysis by Goldstein⁴⁾ shows that the equation for the spinor-spinor system has no discrete eigenvalues also in the case of "ladder approximation" and the total energy $P=0$ (the binding energy equal to the total mass), because of the strong singularity in the interaction kernel.

It is of some interest to study the equation for the spinor-scalar system the singularity of which lies in the middle of the above two cases. We follow the procedure by Wick and take "ladder approximation" with the total energy $P=0$. The equation for our system coupled through the scalar photon field is

$$L\phi(p) = (\gamma_\mu p_\mu - i\kappa) (p^2 + \kappa^2) \phi(p) + \frac{i\lambda}{\pi^2} \int \frac{\phi(q)}{(p-q)^2} [dq] = 0, \quad (1)$$

where $[dq] = dq_1 dq_2 dq_3 dq_4$ is real; ϕ , a spinor function; λ the coupling constant (the eigenvalue of the equation) and p, q are the relative energy-momentum four vectors. For simplicity, we put the masses of both particles equal to κ .

In order to separate the angular variables, let us introduce the following operators;

$$J_\mu = \left(i \frac{\partial}{\partial p_\mu} p_\nu - i \frac{\partial}{\partial p_\nu} p_\mu \right) + \frac{1}{4i} (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu), \quad (2)$$

$$\left. \begin{aligned} K^2 &= \frac{1}{2} \sum_{\mu, \nu} J_{\mu\nu} J_{\mu\nu}, \quad (\mu, \nu = 1, 2, 3, 4) \\ M^2 &= \frac{1}{2} \sum_{i, k} J_{ik} J_{ik}, \quad (i, k = 1, 2, 3) \\ M_3 &= J_{12} \end{aligned} \right\} \quad (3)$$

All these three operators and the operator L defined in (1) commute with each other, so that there exist common eigenfunctions of the operators M_3 , M^2 and K^2 , and ϕ is represented by the linear combination of them. Possible eigenfunctions are, in terms of the polar coordinates $p, \beta, \theta, \varphi$;

$$\psi_1 = \begin{pmatrix} (k-j+1) P_{k+1/2, j-1/2}^{(2)}(\cos \beta) \psi_j^m(\theta, \varphi) \\ i P_{k+1/2, j+1/2}^{(2)}(\cos \beta) \varphi_j^m(\theta, \varphi) \end{pmatrix}, \quad (4a)$$

$$\psi_2 = \begin{pmatrix} (k+j+1) P_{k-1/2, j-1/2}^{(2)}(\cos \beta) \psi_j^m(\theta, \varphi) \\ -i P_{k-1/2, j+1/2}^{(2)}(\cos \beta) \varphi_j^m(\theta, \varphi) \end{pmatrix} \quad (4b)$$

where ψ_j^m and φ_j^m are the well-known eigenfunctions of the total angular momentum operator M^2 and its z-component M_3 belonging to the eigenvalues $j(j+1)$ and m , and $P^{(2)}(\cos \beta)$ is the function in the four dimensional spherical harmonics for angular variable β . The eigenvalue of K^2 is $k(k+2) + \frac{1}{4}$, ($k = 3/2, 5/2, \dots$).

If we assume

$$\phi(p) = -i(2k+3)f_1(p)\psi_1 + (2k+1)f_2(p)\psi_2, \quad (5)$$

we can separate the angular variables. Furthermore, according to Lévy⁶⁾ we project the four dimensional momentum space onto the five dimensional sphere by the stereographic projection. Putting

$$p = \kappa \tan \zeta / 2 \quad (6)$$

and

$$f_1(p) = \sum_{n=0}^{\infty} a_n P_{n+\nu, k+1/2}^{(3)}(\cos \zeta) \quad (7)$$

$$f_2(p) = \sum_{n=0}^{\infty} b_n P_{n+\nu, k-1/2}^{(3)}(\cos \zeta)$$

and substituting (6) and (7) in (1), we can easily integrate it by making use of Heck's theorem⁵⁾. Then, putting the coefficients of each $P_{\mu, \nu}^{(3)}$ equal to zero, we get homogeneous equations with respect to a_n , b_n and finally get the following discrete eigenvalues of the equations:

$$\lambda = (\kappa/2)\nu(\nu+1), \quad \nu \geq k + \frac{1}{2} (= 2, 3, \dots). \quad (8)$$

Thus we reach the conclusion that each

state degenerate with m , j and k , and the situation is the same as the scalar-scalar system.²⁾

When $P \neq 0$, the operators M_z , M^2 and K^2 do not commute with L so that the common eigenfunctions do not exist. We can, however, infer that the degeneracy will be removed because of the asymmetry of the equation projected onto a five dimensional sphere.

- 1) E. E. Salpeter and H. A. Bethe, Phys. Rev. **84** (1951), 1232.
- 2) G. C. Wick, Phys. Rev. **96** (1954), 1124.
- 3) R. E. Cutkosky, Phys. Rev. **96** (1954), 1135.
- 4) J. S. Goldstein, Phys. Rev. **91** (1953), 1516.
- 5) W. Magnus and F. Oberhettinger, *Formeln und Sätze für die Speziellen Funktionen der Mathematischen Physik* (Springer-Verlag 1948).
- 6) M. Lévy, Proc. Roy. Soc. (London) **A 204** (1950), 145.

Constraint in a Quantum System

Yoiti Watanabe

Physics Department, Osaka University, Osaka

September 14, 1956

Let us consider a system of atomic nucleus with Hamiltonian

$$H_1 = (1/2m) \sum_{i=1}^A \mathbf{p}_i^2 + \sum_{i=1}^A V_1(\mathbf{x}_i - \mathbf{R}). \quad (1)$$

We assume a uniform, single-particle nuclear potential about the centre of mass $\mathbf{R} = \sum \mathbf{x}_i / A$. The sum of two-body nuclear interactions $(1/2) \sum_{i,j} V_{i,j}(i, j)$ is here replaced by the average field. Both of them are translationally invariant. There is no approximation between them when they are

of the type of harmonic oscillator, otherwise some consideration would be necessary. In this note, however, these discussions will be set aside though they are of a certain practical interest.

The total momentum $\sum \mathbf{p}_i$, a constant of motion, may be put to vanish

$$\sum \mathbf{p}_i = 0 \quad (2)$$

by an initial condition so as to fix the c.m. at a point. In the classical theory there is another way to fix the c.m. motion. It is called *constraint* and expressed by

$$\sum d\mathbf{x}_i = 0, \quad (3)$$

an opposed to the *subsidiary condition* (2). (3) is rather stringent than (2) in spite of their same results. (2) is deduced from (3), but the reverse does not hold. In the quantum theory subsidiary condition is usually interpreted as a restriction on the state vector like

$$\sum \mathbf{p}_i \Psi = 0, \quad (2)'$$

which is consistent with the equation of motion, while constraint has hitherto had no such a correspondence.

Now, the constraint (3) has effects also on the operation of displacement of positions. In as much as (3) is taken into account, the displacement operator $\partial/\partial \mathbf{x}_i$ is replaced by a form

$$(\partial/\partial \mathbf{x}_i - (1/A) \sum \partial/\partial \mathbf{x}_j). \quad (4)$$

On this standpoint the momentum operators are also modified as

$$\mathbf{p}_i = -i\hbar (\nabla_i - (1/A) \sum \nabla_j), \quad (4)'$$

and then the modified Schroedinger equation

$$H_2 \psi'_n = \left\{ -(\hbar^2/2m) \sum_i (\nabla_i - \sum_j \nabla_j / A)^2 + \sum V(\mathbf{x}_i) \right\} \psi'_n = E_n \psi'_n \quad (5)$$

with the condition $\mathbf{R}\phi'_n=0$, has the similar solutions and eigenvalues to those of

$$H_1\phi_n=E_n\phi_n \quad (6)$$

with the condition (2)', except the δ -factor: that is to say

$$\phi'_n=\phi_n(\mathbf{x}-\mathbf{R})\delta(\mathbf{R}). \quad (7)$$

Conventionally we have the shell-model equation instead of (6)

$$\begin{aligned} H_s\Psi_\alpha &= \left\{ -(\hbar^2/2m)\sum_i \nabla_i^2 + \sum_i V(\mathbf{x}_i) \right\} \Psi_\alpha \\ &= E_\alpha \Psi_\alpha, \end{aligned} \quad (8)$$

for it yields easily so-called Hartree solutions. (8) involves, however, spurious c.m. motions. They should be eliminated. In this case, the condition of the type of (2)' is not available owing to translational asymmetry of H_s , but by the use of (3) Eq. (5) is obtained. Thus our aim is performed. Relation between the set $\{\Psi_\alpha\}$ and $\{\phi'_n\}$, or $\{\phi_n\}$, is given by

$$\phi'_n = \sum_\alpha C_{n\alpha} \Psi_\alpha, \quad (9)$$

$$\sum \{ (E_n - E_\alpha) \delta_{\alpha\beta} + T_{\alpha\beta} \} C_{n\beta} = 0; \quad (10)$$

$$\sum_\alpha \sum_\beta C_{n\alpha}^* \eta_{\alpha\beta} C_{m\beta} = \delta_{nm}. \quad (11)$$

Accordingly the secular equation

$$\det |\delta_{\alpha\beta} (E - E_\alpha) + T_{\alpha\beta}| = 0 \quad (12)$$

is obtained, where

$$T_{\alpha\beta} = -(\hbar^2/2Am) \int \dots \int (\Psi_\alpha^* (\sum_i \nabla_i^2) \Psi_\beta) \cdot d\mathbf{x}_1 \dots d\mathbf{x}_A, \quad (13)$$

$$\eta_{\alpha\beta} = \int \dots \int (\Psi_\alpha^*(\mathbf{x}) \int d\mathbf{R}' \Psi_\beta(\mathbf{x}-\mathbf{R}+\mathbf{R}')) \cdot d\mathbf{x}_1 \dots d\mathbf{x}_A. \quad (14)$$

For light nuclei it is deemed important to mix the configurations of shell-model upon this standpoint. This is pointed out by Elliot and Skyrme, and the result of our

theory agrees with that by them in the same circumstances.¹⁾

In general a number of constraints of Pfaff's form

$$d\chi_\sigma \equiv \sum_{n=1}^N a_\sigma^{(n)}(q) dq_n = 0, \quad \sigma = 1 \dots r < N \quad (15)$$

are taken into account by modifying the commutation relation after the manners of Dirac's generalized Poisson bracket.²⁾ Corresponding to χ_σ , let us define suitable linear forms of conjugate momenta p_n of q_n

$$\phi_\sigma(p, q) \equiv \sum_{n=1}^N \tilde{a}_\sigma^{(n)}(q) p_n, \quad (16)$$

where \tilde{a} 's are chosen so as to satisfy

$$\det |I'_{\sigma\rho}| = \det |\sum_n \tilde{a}_\sigma^{(n)}(q) a_\rho^{(n)}(q)| \neq 0 \quad (17)$$

and to be consistent with the given dynamical terms, viz. the Hamiltonian. Then the canonical commutation relation is modified to

$$\begin{aligned} [p_n, q_m]^* &= -i\hbar \partial_{nm} + i\hbar a_n^{(n)}(q) I^{\sigma\rho} \tilde{a}_\rho^{(n)}(q), \\ I &= (I')^{-1}. \end{aligned} \quad (18)$$

Or, in general, the commutation relation for any variables ξ, η is modified to

$$\begin{aligned} [\xi, \eta]^* &= [\xi, \eta] + (i/\hbar) [\xi, \phi_\sigma] I^{\sigma\rho} [\chi_\rho, \eta] \\ &\quad - (i/\hbar) [\xi, \chi_\sigma] I^{\sigma\rho} [\phi_\rho, \eta]. \end{aligned} \quad (19)$$

$[\xi, \chi_\sigma]$ means $\sum (\partial \xi / \partial q_n) a_\sigma^{(n)}$. (19) satisfies the same algebra as the usual commutator. Particularly (3) yields

$$\begin{aligned} [p_{\mu i}, x_{\nu j}]^* &= -i\hbar \partial_{\mu\nu} (\partial_{ij} - 1/A), \\ \mu, \nu &= 1, 2, 3 \end{aligned} \quad (20)$$

which automatically gives (4).

We can now quantize the internal system in the theory of collective description in which a number of constraints, holonomic and non-holonomic, appeared to eliminate the redundant degrees of freedom derived

by the use of auxiliary variables.³⁾ More generally the present theory is applicable to systems with any constraint. (18) shows that part of dynamical variables would be treated as if they were classical numbers. Such a procedure, dequantization, seems inconsistent with the fundamental concept of quantum theory but in the cases that the suppressed variables are redundant (1) as those introduced owing to a certain postulate of invariance, or (2) as those accidentally mixed on account of a conventional model. Such variables possess no physical significances and should be essentially suppressed. The example of c.m. motion meets with the latter category, while that of internal motion on collective description meets with the former.

However, the procedure may not be so unique. For instance, we have known the free electromagnetic field whose two components are redundant owing to gauge condition. One of them is suppressed by the Lorentz condition

$$\partial A_0 / \partial t - \text{div } \mathbf{A} = 0, \quad (21)$$

and another by one of Maxwell's equations

$$\text{div } \mathbf{D} = 0. \quad (22)$$

On our standpoint (21) corresponds to constraint. It is taken into account by means of modified C. R., for example such as gives

$$[A_0(\mathbf{x}, t) D_0(\mathbf{x}', t)]^* = 0 \quad (23)$$

as a result instead of $[A_0(\mathbf{x}, t), D_0(\mathbf{x}', t)] = i\hbar\delta(\mathbf{x} - \mathbf{x}')$. The old-fashioned Q.E.D. involves it under tacit understanding. Then the straightforward canonical theory shows that corresponding to (22)

$$\text{div } \mathbf{D}^F = 0 \quad (24)$$

becomes then a subsidiary condition independent of (21). These considerations are never important in the Maxwell field but in other kinds of field (e.g. Born-Infeld & c.).

- 1) J. P. Elliot and T. H. R. Skyrme, *Proc. Roy. Soc. A* **232** (1955), 561.
- 2) Y. Watanabe, *Prog. Theor. Phys.* **16** (1956), 1. And Cf. its Reference.
- 3) Y. Watanabe, *loc. cit.*

On an Extended Framework for the Description of Elementary Particles

Mineo IKEDA and Yoshihiko MIYACHI

*Research Institute for Theoretical Physics, Hiroshima University
Takehara-machi, Hiroshima-ken*

(Received July 23, 1956)

An attempt is put forward to formulate as exactly as possible the basic assumption proposed by Pais in his ω -space theory, and a wider framework for the description of elementary particles is obtained which involves that of Pais as a special case. It is found necessary to introduce some field which characterizes the structure of the total manifold. The b field equations introduced by Yang and Mills give one of the admissible forms of the field equations for this new field. Finally the b field is put in comparison with the electromagnetic and the gravitational fields, and similarities and dissimilarities between them are pointed out.

§ 1. Introduction

It is shown from scattering experiments at low energies and from the study of mirror nuclei that the nuclear forces between any two nucleons are very nearly equal. This has led to the charge independence hypothesis of nuclear force and there has so far been no evidence against this hypothesis. Mathematically this hypothesis is expressed as an invariance for rotations in the isotopic spin space. It seems also highly probable that π -mesons contribute to the nuclear forces, hence it is quite natural to assume that the nucleon-meson interaction is also charge independent. We know that recent experiments on π -nucleon scattering have confirmed this assumption.

Hitherto the introduction of the isotopic spin has been done in a somewhat conventional and formal way. Recently Pais¹⁾ has noticed that, in order to intrinsically incorporate the isotopic spin into the theory, the customary description of fields within the framework of space-time is too restrictive. Thus he considered an extension of the framework and made the following basic assumption:

"An element of space-time is not a point but a manifold (" ω -space") which is carried into itself by transformations of the 3-dimensional real orthogonal group."

This ω -space corresponds to the isotopic spin space in the customary theory. It may be expressed as a 2-dimensional sphere embedded in a 3-dimensional Euclidean space. But in the definition of ω -space the group plays an essential role, while any geometrical notion has little meaning. This line of thought is full of interest because it may yield a method of incorporating the isotopic spin in the wave equation of nucleon. Moreover, it seems to present one of the possible forms of the framework suitable to such an attempt to classify a number of elementary particles and to describe their productions, decays and

other reactions in a unified manner.

As a model of possible frameworks founded on the above basic assumption he adopted the following one:*

- P. 1. The total manifold is a direct product of space-time and ω -space.
- P. 2. The group of the total manifold is a direct product of the full Lorentz group and the group in ω -space.

Yet, as was noted by Pais himself, this framework is of a tentative character, and does not seem to be a perfect realization of the above basic assumption. It will be of great interest and of much importance to investigate the possible framework which closely follows the original idea of Pais.

The main part of this paper is concerned with an attempt to formulate his basic assumption as exactly as possible. In the resulting framework, the total manifold and its group are wider than P.1 and P.2 respectively, involving them as special cases. On enlarging the framework, we find it necessary to introduce some field which characterizes the structure of the manifold in a somewhat similar way as the gravitational field does the geometrical structure of space-time. This field is described by the quantities that transform like b_μ introduced by Yang and Mills,²⁾ which fact shows that we may possibly admit the b field as our new field. If, on the other hand, we limit the transformations in the manifold only to P.2, it is equally possible to admit the field equations of different type (e.g., vector meson type equations with isotopic spin 1) for the new field.

It will be interesting to investigate the properties of the new field thus introduced also from another point of view. A number of new particles have been discovered in the last few years and various attempts have been made to give a unified description of their properties. Under such circumstances, it is very likely that we may adopt our new field as one of the models for a Bose particle with spin 1 and isotopic spin 1. The remaining part of this paper is devoted to the investigation on properties of the new field especially of YM type, the similarities and dissimilarities between the b field and the electromagnetic and the gravitational fields being pointed out.

§ 2. Mathematical formulation of the basic assumption

The basic assumption proposed by Pais states that an element of space-time is not a point but a manifold (" ω -space") which is carried into itself by transformations of the 3-dimensional real orthogonal group. In order to distinguish this concept from the customary one, we denote by P an element of space-time regarded as a point, while denoting the one regarded as an ω -space by ω_P . The elements of space-time must be of the same number, i. e., ∞^4 in either case. If two elements of space-time are different from each other when regarded as points, they should be different also as ω -spaces. The index P

* By the word "framework" we mean the total manifold in which the group of coordinate transformations is defined. See reference 1, p. 873.

in ω_P indicates this circumstance, which tacitly implies a one-to-one correspondence between P 's and ω_P 's.

Now we denote by G_P a realization of the 3-dimensional real orthogonal group by means of transformations in ω_P . Its irreducible representation of degree 3 is faithful, and in this section we confine ourselves only to this representation. Denote a basis in the representation space under consideration by E_i^P ($i=1, 2, 3$), and an arbitrary element in it is expressed in the form

$$E^P = E_i^P u^i(P), \quad (2.1)$$

where $u^i(P)$'s, the coordinates of E^P , are dependent on P . In particular, the transforms $'E_i^P$ of E_i^P by a transformation of G_P can be written as

$$'E_i^P = E_j^P a_i^j(P). \quad (2.2)$$

This is a representation of G_P and under a suitable choice of the basis E_i^P the coefficients a_j^i satisfy the equations

$$a_j^i(P) a_k^j(P) = \delta_{jk}. \quad (2.3)$$

Eqs. (2.2) give a transformation of the basis in the representation space, which can also be written in terms of coordinates as

$$u^i(P) = a_j^i(P) 'u^j(P). \quad (2.4)$$

Since ω_P is a manifold which is carried into itself by the 3-dimensional real orthogonal group, it can be represented as an invariant manifold in the representation space. As is well known, such a manifold is given by

$$u^i(P) u^i(P) = R(P)^2. \quad (2.5)$$

The right-hand side can arbitrarily be chosen for a fixed P , so there exist ∞^1 invariant manifolds. In order to get the one-to-one correspondence between P and ω_P we take $R(P)$ to be a one-valued function of P . If we adopt $R(P)E_i^P$ as a new basis,* the right-hand side of (2.5) can be taken to be unity, so that

$$R(P) = 1. ** \quad (2.6)$$

Then E_i^P can be interpreted as forming a basis in ω_P itself, and $u^i(P)$ are coordinates of the element of ω_P , of which only two are independent since there exists a relation (2.5) with (2.6).

So far each of ∞^4 ω -spaces stand quite separately, having no relationship with each other. The framework is too wide as it stands to have any interesting results worthy of special mention. Here we try to establish some connection between them which otherwise remain quite independent. For this purpose we shall assume that ω_{P+dP} can

* We assume $R(P) \neq 0$, otherwise ω_P reduces to an ordinary space-time point P .

** The assumption (2.6) is concerned with the choice of the basis in the representation space and does not refer to the structure of ω -space. We may do without this assumption, when (2.7) becomes an infinitesimal conformal transformation.

be superposed on ω_i so that every element of the former may lie upon the latter. Then the basic elements $E_i^P + dE_i^P$ in ω_{i+1} can be expressed linearly in terms of E_i^P in ω_P . Since the coefficients of the expressions must depend on dP as well as on P , we assume

$$E_i^P + dE_i^P = E_i^P + E_j^P \Omega_i^j(P, dP), \quad \text{i. e.,} \quad dE_i^P = E_j^P \Omega_i^j(P, dP), \quad (2.7)$$

where $\Omega_i^j(P, dP)$ are Pfaffian forms of the coordinates $x^\lambda (\lambda=1, \dots, 4)^*$ of P , i. e.,

$$\Omega_i^j(P, dP) = \Omega_{j\lambda}^i(x) dx^\lambda. \quad (2.8)$$

By using (2.7) an arbitrary element in ω_{P+dP} can easily be written down to be

$$E_i^P u^i(P) + d[E_i^P u^i(P)] = E_i^P [u^i(P) + du^i(P) + \Omega_j^i(P, dP) u^j(P)]. \quad (2.9)$$

The expressions in the bracket on the right-hand side satisfy (2.5) with (2.6). Neglecting infinitesimal quantities of second and higher orders and taking it into account that $u^i(P)$ are arbitrary, we obtain

$$\Omega_j^i(P, dP) + \Omega_i^j(P, dP) = 0.$$

This implies that the infinitesimal transformation (2.7) is an orthogonal transformation.

By (2.7) we can define a one-to-one correspondence between elements of ω_P and of ω_{P+dP} . It should be borne in mind, however, that this can be done only between ω_P and ω_Q when P and Q are at an infinitesimal distance. For two points at a finite distance, the correspondence is determined by specifying some curve which connects P and Q and by superposing ω -spaces along this curve successively. The result may depend in general on the choice of the curve. We can prove that a necessary and sufficient condition for it to be independent of the curve is given by (2.11).

The quantities $\Omega_{j\lambda}^i$ do not appear in the original theory of Pais, so we shall look first into the case in which $\Omega_{j\lambda}^i$ identically vanish in some detail. In this case it is clear that $dE_i^P = 0$ for every P . Therefore we can choose the basis in ω -space quite independently of the coordinates x^λ of P . This fact results in that a point in the total manifold is specified by x^λ and the coordinates u^i (with (2.5) and (2.6)) in one particular ω -space, and the total manifold can be considered as reducing to P. 1. The conditions $\Omega_{j\lambda}^i = 0$, however, are not invariant under the transformation (2.4) in ω -space, as can readily be seen from

$${}^i\Omega_{j\lambda}^i = \bar{a}_k^i \Omega_{i\lambda}^k a_j^k + \bar{a}_k^i \partial_\lambda a_j^k, \quad \bar{a}_j^i a_k^j = \delta_k^i. \quad (2.10)$$

This has been obtained from (2.7) in the primed system by using (2.2) and (2.7) and by taking into account the fact that dx^λ are quite arbitrary. If, however, we restrict the transformations in the manifold to P. 2, then a_j^P s do not depend on x^λ and the above conditions remain valid after the transformation. In the work of Pais, the group of the total manifold is restricted only to such a case. Thus the framework originally assumed by Pais is characterized by $\Omega_{j\lambda}^i = 0$.

* Throughout this paper Greek indices stand for the values 1, ..., 4 corresponding to the space-time coordinates, while Latin ones 1, 2, 3 corresponding to the coordinates in ω -space.

Finally, we look for a condition that the given $\mathcal{Q}_{j\lambda}^i$ be reducible to zero by a suitable transformation of the basis in ω -space. For this it is necessary and sufficient that there exist a_j^i for which the right-hand side of (2.10) vanishes. The conditions of integrability turn out to be

$$\mathcal{Q}_{j\lambda\mu}^i = 0, \quad (2.11)$$

where

$$\mathcal{Q}_{j\lambda\mu}^i = \partial_\mu \mathcal{Q}_{j\lambda}^i - \partial_\lambda \mathcal{Q}_{j\mu}^i + \mathcal{Q}_{k\mu}^i \mathcal{Q}_{j\lambda}^k - \mathcal{Q}_{k\lambda}^i \mathcal{Q}_{j\mu}^k. \quad (2.12)$$

We can easily ascertain that (2.11) are invariant with respect to both (2.4) and the Lorentz transformations.

It need hardly be stressed that, from the standpoint of the geometry of connections, the results in this section can be interpreted in the following manner: ω -space is associated with every point in space-time regarded as a basic space, and $\mathcal{Q}_{j\lambda}^i$ and $\mathcal{Q}_{j\lambda\mu}^i$ can be identified with the coefficients of the connection and the curvature tensor, respectively.³⁾

§ 3. Interpretation of $\mathcal{Q}_{j\lambda}^i$

The quantities $\mathcal{Q}_{j\lambda}^i$ did not appear in the original theory of Pais, while they have played an essential role in enlarging the framework of description. In order to see the structure of the enlarged framework more clearly, it seems necessary to make a further examination of these quantities.

Now, of any theory in physics, we take it for granted that if two sets of quantities are connected by a transformation admitted in that theory, they should describe the same physical object and are physically equivalent even though they have different expressions. Therefore, $\mathcal{Q}_{j\lambda}^i$, for which (2.11) holds good, can be considered to have no physical meaning, since we are able to make them vanish by the transformation (2.4). On the other hand, for $\mathcal{Q}_{j\lambda}^i$, for which (2.11) is not true, we fail to make them vanish whatever basis we may choose in ω -space. Consequently, such $\mathcal{Q}_{j\lambda}^i$ must no longer be merely auxiliary quantities, but need to be regarded as quantities of some physical meaning. Below we shall study this case in some detail.

In the isotopic spin formalism every kind of particles is associated with some of irreducible representations of G_P . Here we consider the one associated with a representation of degree n . If we denote its wave function by ψ , then the transformation (2.4) of G_P is represented by

$$\psi = S(x)' \psi, \quad (3.1)$$

where $S(x)$ is an $n \times n$ matrix and depends on x^λ in general. Similarly, (2.7) can also be represented by an $n \times n$ matrix, which is shown to have a quite simple expression as follows.

Since (2.7) is an infinitesimal orthogonal transformation, it can be expressed linearly in terms of three independent infinitesimal transformations. Therefore, if we write $\mathcal{Q}_{j\lambda}^i$ in a matrix form \mathcal{Q} , rows and columns being indicated by i and j respectively, we have

$$\Omega(P, dP) = \omega^i(P, dP) O_i, \quad (3.2)$$

where O_i are the matrices corresponding to three independent infinitesimal transformations of G_P . They can be chosen independently of both P and dP ; e. g.,

$$O_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad O_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad O_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Moreover, we find by (2.8) that ω^i are Pfaffian forms of x^λ , i. e., of the form

$$\omega^i(P, dP) = \omega_\lambda^i(x) dx^\lambda. \quad (3.3)$$

Now, in the representation under consideration the matrices O_i are represented by $n \times n$ matrices I_i , which satisfy the same relations as O_i , i. e., $\epsilon^{ijk} I_j I_k = I_i$. Here ϵ^{ijk} 's mean permutation symbols. Then the matrix (3.2) is represented by $\omega^i I_i$. Eventually, of n^2 elements of the matrix ω which represents (2.7), only three are independent regardless of the value n . Moreover, the coefficients ω^i by which the matrix ω is expressed linearly in terms of I_i are common to every n .

As has been mentioned previously, it seems quite reasonable to consider that Ω_{jk}^i 's which do not satisfy (2.11) will represent something of physical reality, which we shall assume here to be some field. Then, in the presence of this field the total manifold is of a type different from P. 1, while it reduces to P. 1 in the absence of this field. Therefore we can assert that this field characterizes the structure of the total manifold.* This rather interesting conclusion reminds us of the fact that the gravitational field characterizes the geometrical structure of space-time. From our foregoing arguments, this field is described by twelve ω_λ^i . Since they are common to every n , this field can be considered to interact with particles which belong to any irreducible representation.

The transformation (2.4) in ω -space or its representation (3.1) corresponds to the isotopic gauge transformation in YM. We can see that the matrix field B_λ in YM transforms under isotopic gauge transformation just in the same way as do the matrices $I'_\lambda = (i/\epsilon) \omega_\lambda^i I_i$ under (2.4). In fact, we have from (2.10)

$$I'_\lambda = S^{-1} I'_\lambda S + (i/\epsilon) S^{-1} \partial_\lambda S,$$

by replacing a_j^i and Ω_{jk}^i by the matrices S and $-i\epsilon I'_j$, respectively. This is nothing but the one obtained in YM for B_λ . The similarities between I'_λ and B_λ can also be seen as follows. Consider an arbitrary element (2.9) in $\omega_{P \rightarrow dP}$. Replacing u^i and Ω_{jk}^i by the wave function ψ^i and by $-i\epsilon I'_\lambda$, respectively, we find in (2.9) the following expressions:

$$\partial_\lambda \psi^i - i\epsilon I'_\lambda \psi^i.$$

The appearance of derivatives of ψ^i in such a combination is just what was assumed at

* Notice that no specific assumption is made concerning the structure of individual ω -space.

the beginning of YM. Since $(i/2\epsilon)\omega_\lambda^{ij}$ s are to Γ'_λ what in YM b_λ^{ij} s are to B_λ , we have the possibility of admitting field equations for the b field in YM as those for our new field. Therefore, our framework can be said to contain the framework adequate to the YM theory.

Nevertheless, we have still another possibility of admitting field equations of different type. As was mentioned above, our new field characterizes the structure of the total manifold. Consequently, even if we confine ourselves only to those transformations of P. 2 as tentatively assumed by Pais, the introduction of the new field is still necessary in general in our enlarged framework. Under such a restriction, field equations of linear type can also be admitted for the new field (e.g., equations of the vector meson type for the field with isotopic spin 1), whereas they cannot in YM.

§ 4. Properties of the field of YM type

As has been shown in the foregoing sections, it is necessary in general to introduce a new Bose field in our enlarged framework.

Meanwhile a number of new particles have been discovered in recent years and many attempts have been proposed to give them a unified description. In view of these circumstances it is very likely that this new field may be adopted as a possible model of these new particles. As was mentioned in the last section the b field equations in YM can be admitted for the new field. Since properties of the field of such a type have hardly been clarified up to now, we study them hereafter.

The field equations of the b field interacting with the nucleon field are, after YM,²⁾

$$\partial_\nu f_{\mu\nu}^i + 2\epsilon\epsilon^{ijk} b_\nu^j f_{\mu\nu}^k + J_\mu^i = 0, \quad (4.1)$$

$$\gamma_\mu (\partial_\mu - i\epsilon\tau_i b_\mu^i) \psi + m\psi = 0 \quad (4.2)$$

where

$$J_\mu^i = i\epsilon\bar{\psi}\gamma_\mu\tau_i\psi, \quad (4.3)$$

$$f_{\mu\nu}^i = \partial_\nu b_\mu^i - \partial_\mu b_\nu^i - 2\epsilon\epsilon^{ijk} b_\mu^j b_\nu^k. \quad (4.4)$$

Further, it is assumed that

$$\partial_\mu b_\mu^i = 0. \quad (4.5)$$

By using (4.5), (4.1) can be written as

$$\begin{aligned} \square b_\mu^i &= -J_\mu^i - f_\mu^i \\ &\equiv -J_\mu^i - 4\epsilon\epsilon^{ijk} b_\nu^j \cdot \partial_\nu b_\mu^k - 2\epsilon\epsilon^{ijk} \partial_\mu b_\nu^j \cdot b_\nu^k + 4\epsilon^2 (b_\mu^i b_\nu^j b_\nu^j - b_\nu^i b_\mu^j b_\nu^j). \end{aligned} \quad (4.6)$$

On the other hand, as is well known, the electromagnetic field is described by the field equations

$$\square A_\mu = -j_\mu, \quad (4.7)$$

with the supplementary condition

$$\partial_\mu A_\mu = 0, \quad (4.8)$$

where j_μ is the current 4-vector. Also it was shown by Gupta¹ that the field equations of Einstein's gravitational field

$$R^{\mu\nu} - g^{\mu\nu} R/2 = -\kappa^2 T^{\mu\nu}$$

can be written in an alternative form as follows:

$$\square g^{\mu\nu} = -\kappa^2 \theta^{\mu\nu} = -\kappa^2 (T^{\mu\nu} + t^{\mu\nu}) \quad (4.9)$$

with the coordinate conditions

$$\partial_\nu g^{\mu\nu} = 0, \quad (4.10)$$

where $T^{\mu\nu}$ and $t^{\mu\nu}$ are the energy-momentum tensor densities of the matter and of the gravitational fields, respectively.*

If we may regard the right-hand side of (4.6) as some physical quantity of the total system, then we are able to say that (4.5) and (4.6) have a remarkable similarity to (4.8) and (4.7) and to (4.10) and (4.9). In fact, we shall show in what follows that the expression $J_\mu^i + j_\mu^i$ in (4.6) can be defined as the isotopic spin current density of the total system. For this we use a method similar to that of Gupta¹ by which he obtained the Lagrangian density in flat space for the gravitational field. This method may seem at first to be a detour to our purpose. Nevertheless, similarities and dissimilarities between the electromagnetic, the gravitational and the b fields will be seen more clearly by this method.

We start from the following Lagrangian density:

$$\mathcal{L}_0 = -1/2 \cdot \partial_\nu b_\mu^i \cdot \partial_\nu b_\mu^i + \partial_\mu b_\nu^i \cdot \partial_\nu b_\mu^i - \bar{\psi} (\gamma_\mu \partial_\mu + m) \psi. \quad (4.11)$$

Using (4.5), we can see that the field equations resulting from \mathcal{L}_0 are

$$\square b_\mu^i = 0, \quad (4.12)$$

$$(\gamma_\mu \partial_\mu + m) \psi = 0. \quad (4.13)$$

Now, we define, after Bludman,⁽¹⁾ the isotopic spin current density \mathcal{N}_μ^i of the system by

$$\mathcal{N}_\mu^i = -i\epsilon \left[(\partial_\mu \bar{\psi} \cdot \partial_\nu \psi) \tau_i^{\nu\mu} - \bar{\psi} \tau_i^{\nu\mu} (\partial_\mu \psi) \right] + 2\epsilon^{\nu\mu k} b_\mu^j (\partial_\nu / \partial b_{\nu,\mu}^k), \quad (4.14)$$

where $\psi_{,\mu} = \partial_\mu \psi$ and $b_{\nu,\mu}^i = \partial_\mu b_\nu^i$. Then \mathcal{L}_0 gives the isotopic spin current density

$$i\epsilon \bar{\psi} \gamma_\mu \tau_i \psi + 4\epsilon^{\nu\mu k} b_\nu^j \cdot \partial_\nu b_\mu^k - 2\epsilon^{\nu\mu k} b_\nu^j \cdot \partial_\mu b_\nu^k \equiv J_\mu^i + j_\mu^{(0)i}. \quad (4.15)$$

We now have to modify \mathcal{L}_0 in such a way that in the resulting equation the quantity (4.15) appears on the right-hand side of (4.12). For this we take

$$\mathcal{L}_1 = \mathcal{L}_0 + \mathcal{L}', \quad (4.16)$$

* $t^{\mu\nu}$ is not the energy-momentum pseudo-tensor density of the gravitational field with which we are familiar in general relativity, but is the symmetrical one which is obtained through Belinfante's method.⁽⁴⁾⁽⁵⁾

$$\mathfrak{L}' = J_\mu^i b_\mu^i + 2\epsilon \epsilon^{ijk} \partial_\nu b_\mu^i \cdot b_\mu^j b_\nu^k. \quad (4.17)$$

However, then \mathfrak{L}_1 gives

$$J_\mu^i + j_\mu^{(0)i} - 4\epsilon^2 (b_\mu^i b_\nu^j b_\nu^j - b_\nu^i b_\mu^j b_\nu^j) \equiv J_\mu^i + j_\mu^{(0)i} + j_\mu^{(1)i}. \quad (4.18)$$

Again we have to modify \mathfrak{L}_1 in such a way that in the resulting equation the quantity (4.18) appears on the right-hand side of (4.12). For this we choose a Lagrangian of the form

$$\mathfrak{L}_2 = \mathfrak{L}_0 + \mathfrak{L}' + \mathfrak{L}'', \quad (4.19)$$

$$\mathfrak{L}'' = -\epsilon^2 (b_\mu^i b_\mu^i b_\nu^j b_\nu^j - b_\mu^i b_\nu^j b_\mu^j b_\nu^j). \quad (4.20)$$

Since \mathfrak{L}'' in \mathfrak{L}_2 does not contain $b_{\lambda,\mu}^i$, we find from the definition of \mathfrak{S}_μ^i that the isotopic spin current density derived from \mathfrak{L}_2 is the same as that derived from \mathfrak{L}_1 . At the same time, the field equations resulting from \mathfrak{L}_2 are easily seen to be just (4.2) and (4.6). Therefore our procedure comes to an end here. After all, \mathfrak{L}_2 gives the Lagrangian density of the total system: the resulting field equations turn out to be (4.2) and (4.6), and the right-hand side of (4.6) can be interpreted as the isotopic spin current density of that system.* Consequently, the formal similarity mentioned above is established.

It should be remembered here that although there is such a similarity among these three fields, there are also some dissimilarities between them. The above derivation of the Lagrangian for the b field shows that the non-linear character of the b field equations is due to the fact that the b field has its own isotopic spin current density, which can again serve as a source of the field. This situation is quite similar to that in the gravitational field. Indeed, it was shown by Gupta⁽⁵⁾ that the non-linear character of the gravitational field equations can be attributed to the fact that the energy-momentum tensor of the gravitational field acts as a source. This is different from the case of the electromagnetic field, which is itself chargeless, and consequently satisfies linear equations in the absence of a charged field.

Furthermore, we have shown that our procedure to get the Lagrangian density for the b field comes to an end after a few steps. On the contrary, for the gravitational field the Lagrangian density in flat space consists of infinite series of terms, as has been shown by Gupta.⁽⁵⁾ This is because $t^{\mu\nu}$ are of the same degree as \mathfrak{L} with respect to derivatives of the field variables, whereas j_μ or \mathfrak{S}_μ^i are lower than \mathfrak{L} by one. In this respect the b field seems to be more akin to the electromagnetic than to the gravitational field. It is quite interesting to observe that such a specific feature of the individual field comes from the different physical nature of the source (e.g., such as the charge-current, the energy momentum or the isotopic spin current) of each field, rather than from the different values of the spin or the isotopic spin.**

Thus we have seen that there are several similarities between the electromagnetic, the

* \mathfrak{L}_2 differs from the Lagrangian density given in YM only by the second term in \mathfrak{L}_0 .

** Cf. reference 5, p. 1685.

gravitational and the \mathbf{b} fields on the one hand, and that there are also some dissimilarities which come from the character of the source of the individual field. The \mathbf{b} field shares the common properties with the gravitational field in that they both satisfy non-linear equations and that their non-linear parts can be treated analogously. It also shares the common properties with the electromagnetic field in that their Lagrangian densities are given in closed forms.* To sum up, the \mathbf{b} field seems to have a character intermediate between the electromagnetic and the gravitational fields.

§ 5. Concluding remarks

In order to intrinsically incorporate the isotopic spin into the theory, Pais proposed a basic assumption how to widen the customary framework for the description of elementary particles. However, the actual framework devised by him does not seem to be a perfect realization of his original idea. In the present paper we have tried to realize his basic assumption as perfectly as possible. The framework thus obtained involves that of Pais as a special case, and except this case it is necessary to introduce some field, which characterizes the structure of the total manifold in the enlarged framework. It is even possible that this new field be the \mathbf{b} field of YM. If, however, the group of transformations in the manifold is restricted only to P. 2 as tentatively assumed by Pais, then not only the \mathbf{b} field equations in YM but some linear equations can also be admitted for the new field. Finally we have compared the new field of YM type with the electromagnetic and the gravitational fields, and similarities and dissimilarities between them have been pointed out.

The fact that the structure of the total manifold is characterized by a new field is of much interest to the present authors. It is true that the arguments based on the requirement of invariance as was given by Utiyama⁷⁾ provide one way of introducing a new field, but it will be promising as well to consider a new field from the standpoint of the structure of the manifold. Thus it will be of great importance to construct a geometry of the enlarged manifold.

The authors wish to express their sincere thanks to Professor Y. Mimura and Professor H. Takeno for their continual encouragement and helpful discussions throughout this work. They are also grateful to Professor S. Machida for his valuable discussions at the preliminary stage of this work.

* (4.1) with $J_\mu{}^i=0$ can be solved when the field $b_\mu{}^i$ is static and invariant under spatial rotations. Its general solution is given by

$$b_\lambda{}^i=0 \quad \text{for } \lambda=1, 2, 3, \quad b_4{}^i=i(\alpha^i/r+\beta^i),$$

where α^i and β^i are constants which satisfy $\alpha^i\beta^i=0$. (The proof is omitted here.) This also shows another similarity between the \mathbf{b} and the electromagnetic fields.

References

- 1) A. Pais, *Physica* **19** (1953), 869.
- 2) C. N. Yang and R. L. Mills, *Phys. Rev.* **96** (1954), 191. We refer to this paper as YM.
- 3) For example, É. Cartan, *Leçons sur la théories desespaces à connexion projective*, Paris (1937).
- 4) S. N. Gupta, *Proc. Phys. Soc.* **65A** (1952), 608.
- 5) S. N. Gupta, *Phys. Rev.* **96** (1954), 1683.
- 6) S. A. Bludman, *Phys. Rev.* **100** (1955), 372.
- 7) R. Utiyama, *Phys. Rev.* **101** (1956), 1597.

Meson-Nucleon Scattering in Low Energy Region

Daisuke ITÔ and Shigeo MINAMI*

*Physics Institute, Tokyo University of Education, Tokyo
and**Department of Physics, Osaka City University, Osaka**

(Received May 21, 1956)

The scattering length a for the s -wave of the process $\pi^+ + p \rightarrow \pi^+ + p$ is calculated by tracing the experimental data for α_3 . Comparing the value of $4\pi a^2$ with the theoretical value of cross section in the Thomson limit, it is found that the coupling constant $f^2/4\pi$ is nearly equal to 1.

Because of the importance of meson-meson interaction inferred from the analysis of meson reaction at 1.5 Bev, the scattering cross sections by this interaction are estimated and are compared with those by the usual meson-nucleon interaction. When the $(\lambda/16)(\phi_\alpha \phi_\alpha)^2$ -interaction is adopted as a simple model, the effect of this interaction is very small in the case of the value of $(\lambda/4\pi)^2 \cong 1.5$ which has been adopted to explain the $\pi\text{-}2\pi$ process at 1.5 Bev

§ 1. Introduction

At one time, the strange behavior of α_3 in the neighbourhood of 30~50 Mev was reported and discussed from various points of view, but this problem has been resolved by the increasing accuracy of experimental data. It is now important to explain the relative sign of α_3 and α_1 from the standpoint of meson theory. Bosco and Stroffolini,¹⁾ adopting only the interaction of $\pi \rightleftharpoons N + \bar{N}$ and making use of the method employed by Lee,²⁾ have recently shown, with the reasonable cut-off momentum, the possibility of explaining the experimental result. In this case, moreover, the effective coupling constant $f^2/4\pi = 1$ was taken in order to obtain the agreement with the phase-shift analysis given by Orear³⁾. This fact may be interpreted as due to the s -wave damping; however, there may be controversy about that effect from field theoretical point of view, since their method is based on the Tamm-Dancoff approximation and a special choice of type of interaction. But it is not our purpose to give a thorough discussion of this point.

First of all, let us trace the reason why the calculated result obtained by Bosco and Stroffolini agrees fairly well with the experimental result. Which is more essential of the two approximations:

- (i) Only the interaction $\pi \rightleftharpoons N + \bar{N}$ is taken into consideration;
- (ii) the ladder approximation and cut-off theory are used?

A simple examination shows that (i) is more essential. Then we will assume the system with Bosco-Stroffolini's choice of interaction for a model to describe the s -wave scattering without introducing the meson-meson interaction, and roughly estimate the cross section

in the Thomson limit for $\pi^+ + p \rightarrow \pi^+ + p$. On the other hand, the scattering length for the s -wave of this process will be calculated by tracing the experimental data for α_{π} . Making use of these results, it will be found that the coupling constant is $f^2/4\pi \cong 1$.

Although a number of approaches to meson-nucleon scattering have so far been done mainly through the meson-nucleon interaction, we think it worth-while, for an extensive view of meson reaction, to take into account of the effect of meson-meson interaction in low energy region, because of the importance of meson-meson interaction inferred from the analysis of meson reaction at 1.5 Bev⁽⁴⁾⁽⁵⁾⁽⁶⁾.

With regard to π - π interaction, several approaches have been tried from various points of view by many authors,⁽⁷⁾⁽⁸⁾⁽⁹⁾ and, as a simple model, the interaction of $\lambda(\phi_\alpha \phi_\alpha)^2$ has been examined mainly in the process of $\pi-2\pi$ ⁽¹⁰⁾⁽¹¹⁾. With the purpose mentioned above we intend to study the contribution of this $\lambda(\phi_\alpha \phi_\alpha)^2$ -interaction to the meson-nucleon scattering. In § 3, the scattering cross sections by this interaction will be estimated and be compared with those by the usual meson-nucleon interaction.

§ 2. The phase shifts of s -wave and coupling constant

In order to examine the method employed by Bosco and Stroffolini, let us take only the approximation (i) and leave out the approximation (ii) mentioned in § 1. In this case, the simplest way to know the result is to perform the calculation by means of the perturbation theory adopting only the interaction of $\pi \rightleftharpoons N + \bar{N}$.

When the R -matrices of the states $I=3/2$ and $I=1/2$ of the isotopic spin are represented by the notation of $R_{3/2}$ and $R_{1/2}$ respectively, the amplitudes of the meson-nucleon scattering, as is well known, are expressed as follows:

$$\begin{aligned} \pi^+ + p &\rightarrow \pi^+ + p & R_{3/2}, \\ \pi^- + p &\rightarrow \pi^- + p & (1/3) (R_{3/2} + 2R_{1/2}), \\ \pi^- + p &\rightarrow \pi^0 + n & (\sqrt{2}/3) (R_{3/2} - R_{1/2}). \end{aligned} \quad (1)$$

The graphs corresponding to π^+ and π^- meson scattering by proton (except charge exchange) are Fig. 1 and Fig. 2 respectively, but the contribution from Fig. 2 must be

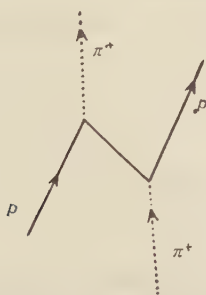


Fig. 1.

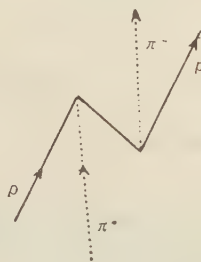


Fig. 2.

neglected so long as the transition through only the interaction of $\pi \rightleftharpoons N + \bar{N}$ is taken into consideration.

Therefore

$$R_{3/2} + 2R_{1/2} = 0. \quad (2)$$

As our attention is here directed to the phenomena in low energy region,

$$R_{3/2} \cong \exp(2i\alpha_3) - 1 \cong 2i\alpha_3,$$

$$R_{1/2} \cong \exp(2i\alpha_1) - 1 \cong 2i\alpha_1.$$

From eq. (2) we see that the sign of α_1 is different from that of α_3 .^{*)} Although the quantitative agreement cannot be obtained, we may infer from the above discussion that the approximation (i) is more essential than (ii).

On this standpoint, let us study the meson-nucleon scattering in low energy region from the phenomenological point of view. Since the transition $\pi^- + p \rightarrow \pi^- + p$ may be neglected and, as the result, the relation between $R_{3/2}$ and $R_{1/2}$ is given by (2), we have only to consider the process of $\pi^+ + p \rightarrow \pi^+ + p$. The description of the π - N scattering phenomena may be more or less different from that of N - N scattering. But if we refer to the idea of the potential giving rise to the π - N scattering as is shown by Sawada¹²⁾, the treatment of the π - N scattering can be performed along the same line as in the case of the N - N scattering.

Thus the s -wave phase shift at low energy can be expressed approximately as

$$k \cot \alpha_3 = -(1/a) + (k^2/2) \cdot r_0. \quad (3) \quad **)$$

We will here adopt the values both of $\alpha_3 = -4.9^\circ$ at 58 Mev obtained by Podansky et al.¹⁴⁾ and of $\alpha_3 \cong -2.5^\circ$ at 23 Mev obtained by Orear¹⁵⁾. Putting these values into eq. (3),

$$\begin{aligned} a &= 1.05 \times 10^{-14} \text{ cm}, \\ r_0 &= 1.48 \times 10^{-12} \text{ cm}. \end{aligned} \quad (4)$$

Then the energy dependence of α_3 becomes as Fig. 3. It fairly agrees with the experimental data up to the energy of about 100 Mev, in spite of the fact that the assumption of shape independence is considered to be effective in lower energy region. We may think that this circumstance is due to the following reason.

In the nuclear scattering, as is well known¹⁶⁾, there is the following relation:

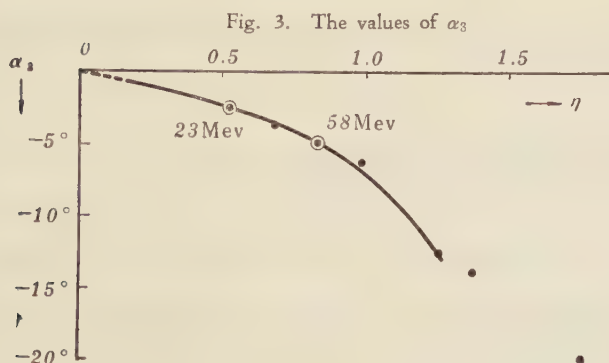
$$k \cot \alpha_3 = -1/a + (k^2/2) \rho(0, E)$$

$$\text{with} \quad (1/2) \rho(0, E) = \int_0^\infty (\psi_0' \psi - u_0 u) dr, \quad (3')$$

where u is the radial wave function multiplied by r and ψ represents the asymptotic be-

*) It is easily found that α_3 is negative.

**) If we follow the formula of Chew and Low,¹³⁾ $k \cot \alpha_3/\omega^* = A + B\omega^*$, the graph shown in Fig. 3 undergoes some change in high energy region. However, the value of coupling constant $f^2/4\pi$ is 1.1, nearly the same as in eq. (10).



havior of u for large distance. Let us consider the expansion of $\rho(0, E)$ in the neighbourhood of $E=E_1$ (in our case $E_1 \simeq 40$ Mev) instead of $E=0$. Then

$$\rho(0, E) = \rho(0, E_1) + \frac{\rho'(0, E_1)}{1!} (E - E_1) + \dots \quad (5)$$

Taking only the first term of eq. (5), we set the following approximation:

$$\rho(0, E) \simeq \rho(0, E_1). \quad (6)$$

Then eq. (6) is not a very good approximation for the energy far apart from $E=E_1$. In the lower energy region (k is of the order of 10^{-12} cm $^{-1}$), however, the roughness of this approximation does not become so serious in examining the amount of $k \cot \alpha_3$, because the first term of the right-hand side in (3) contributes predominantly (10~100 times), rather than the second one multiplied by k^2 .

Moreover, in lower energy region where the second term of eq. (3) may be neglected, α_3 is approximately expressed as follows:

$$\alpha_3 = -0.074 \gamma. \quad (7)^*)$$

This relation is more or less different from that of $\alpha_3 = -0.11 \gamma$ given by Orear³⁾, but is nearly equal to that of $\alpha_3 = -0.083 \gamma$ obtained by Spry or the experiments for mesic atoms¹⁷⁾.

Employing the value of a thus obtained, the cross section for $\pi^+ + p \rightarrow \pi^+ + p$ in the Thomson limit can easily be calculated by

$$\sigma = 4\pi a^2. \quad (8)$$

On the other hand, it is also estimated straightforwardly by the perturbation theory and its amount turns out to be

$$\sigma = 4\pi \frac{(f^2/4\pi)^2}{m^2} \cdot \frac{1}{(1 + \mu/m)^2 (2 - \mu/m)^2}, \quad (9)$$

where m and μ are the masses of nucleon and meson respectively. When we call to mind that the calculation may, as is shown above, give qualitatively correct prediction if

*) From eq. (2), it is easily seen that $\alpha_1 = 0.037 \gamma$.

only the assumption (i) is satisfied, this result may be considered to be probably reliable^{*)}. By putting (8) and (9) equally,

$$f^2/4\pi \cong 1.07. \quad (10)$$

It is very interesting that this value is nearly equal to that of the effective coupling constant $f^2/4\pi=1$ obtained by Bosco and Stroffolini.

§ 3. Contribution of Kovacs term to π - N scattering

Even if the importance of π - π interaction is suggested from the analysis of the meson reaction in Bev region, it does not necessarily follow that the $\lambda(\phi_\alpha \phi_\alpha)^2$ -term (Kovacs term) plays an important role. But it is worth-while to examine this interaction as one of the simplest models. For π - 2π process, the calculation taking this term into account have been done by Kovacs¹⁰⁾ and Ito, Mori and Yamazaki.¹¹⁾ We intend to examine the effect of the Kovacs term in low energy phenomena (π - N scattering). For simplicity, let us estimate the order of the contribution by the lowest order calculation. We think it necessary to compare this value with that obtained by the perturbation theory under the assumption (i) because, in § 2, we have seen that the latter tells us the correct order of magnitude for the cross section.

Assuming the following interaction Hamiltonian

$$H_{int} = if\bar{\psi}\gamma_5\tau_\alpha\phi_\alpha\psi + (\lambda/16)(\phi_\alpha\phi_\alpha)^2, \quad (11)$$

we write down, first of all, the result for $\pi^+ + p \rightarrow \pi^+ + p$. The necessary diagrams are shown in Fig. 4 and Fig. 5, and the differential cross sections in the center-of-mass system are expressed by the following equation,

$$\frac{d\sigma(+)}{d\Omega} = (1+1/4)^2 \frac{(f^2/4\pi)^2 (\lambda/4\pi)^2 (B/2\pi)^2}{(\sqrt{1+p^2/m^2} + \sqrt{\varepsilon + p^2/m^2})^2} \cdot \frac{(1+b)}{8m^2}, \quad (12)$$

$$b = -(IF)/m^2, \quad \varepsilon = \mu^2/m^2,$$

$$B = \int_0^1 dx \int_0^x dy \frac{x-y}{\varepsilon - \varepsilon(x-y) + (x-y)^2 + 2(b-1)y(1-x)}, \quad (13)$$

where I and F represent the four-dimensional momenta of the nucleon in the initial and final states respectively, and p is the magnitude of the meson momentum in center-of-mass

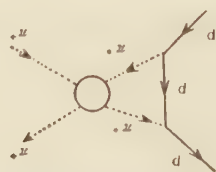


Fig. 4.



Fig. 5

*) It is useful to remember the conclusion deduced by Deser, Thirring and Goldberger¹⁹⁾.

system. The coefficients 1 and 1/4 in the factor $(1+1/4)^2$ in eq. (12) show the contributions from the transition matrices corresponding to Fig. 4 and Fig. 5 respectively. Now,

$$b = - (IF) / m^2 = 1 + p^2 (1 - \cos \theta) / m^2 \geq 1,$$

therefore, the integrand in eq. (13) is definitely positive, and the calculation can easily be carried out. The results are shown in Table 1.

Table 1. The values of $(d\sigma(+)/d\Omega)/(f^2/4\pi)^2(\lambda/4\pi)^2$ in center-of-mass system (10^{-29} cm²/sterad.)

| Angle Energy | 0° | 30° | 60° | 90° | 120° | 150° | 180° |
|-----------------|------|------|------|------|------|------|------|
| 24 Mev | 0.46 | 0.45 | 0.45 | 0.44 | 0.43 | 0.42 | 0.42 |
| 58 Mev | 0.43 | 0.42 | 0.41 | 0.39 | 0.37 | 0.36 | 0.35 |

Table 2. The values of $(c\sigma(+)/d\Omega)/(f^2/4\pi)^2$ in center-of-mass system (10^{-29} cm²/sterad.)

| Angle Energy | 0° | 30° | 60° | 90° | 120° | 150° | 180° |
|-----------------|-----|-----|-----|-----|------|------|------|
| 24 Mcv | 9.2 | 9.2 | 9.3 | 9.5 | 9.7 | 9.8 | 9.9 |
| 58 Mev | 8.5 | 8.6 | 8.9 | 9.2 | 9.6 | 9.9 | 10.1 |

For the convenience of comparison, the results derived from the perturbation theory are also given in Table 2.

It is easily found that the differential cross sections $d\sigma(-)/d\Omega$ for $\pi^- + p \rightarrow \pi^- + p$ are the same with $d\sigma(+)/d\Omega$. For $\pi^- + p \rightarrow \pi^0 + n$, on the other hand, the contributions from Fig. 6 and Fig. 7 come to cancel each other, and the matrix element turns out to be zero.

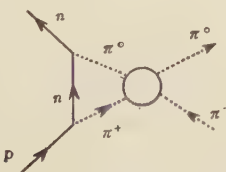


Fig. 6.

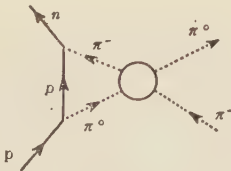


Fig. 7.

Therefore, from eq. (1)

$$(\sqrt{2}/3) (R_{1/2} - R_{1/2}) = 0.$$

This result clearly contradicts with the experimental data. Although any definite conclusion may not be reached through the calculation of the lowest order, it may be found from the above consideration that if the Kovacs term plays an important role the behavior of s-wave in low energy will be hard to explain. This difficulty, however, will make

no problem in the case of the value of $(\lambda/4\pi)^2 \simeq 1.5^{*})$ which has been adopted to explain the $\pi-2\pi$ process at 1.5 Bev. (c.f. Table 1 and Table 2)

In conclusion, the authors express their sincere thanks to Prof. Z. Koba, Prof. Y. Yamaguchi and Dr. H. Hasegawa for their valuable discussions.

References

- 1) B. Fosco and R. Stroffolini, *Nuovo Cimento* **1** (1955), 433.
- 2) T. D. Lee, *Phys. Rev.* **95** (1954), 1329.
- 3) J. Orear, *Phys. Rev.* **100** (1955), 288.
- 4) Eisberg, Fowler, Lea, Shaphard, Shutt, Thorndike and Whittemore, *Phys. Rev.* **97** (1955), 797.
- 5) W. D. Walker and J. Crussard, *Phys. Rev.* **98** (1955), 1416.
- 6) D. Ito and S. Minami, *Prog. Theor. Phys.* **14** (1955), 482.
- 7) G. Takeda, *Phys. Rev.* **100** (1955), 440.
- 8) F. J. Dyson, *Phys. Rev.* **99** (1955), 1037.
- 9) Z. Koba, *Prog. Theor. Phys.* **15** (1956), 294.
- 10) J. S. Kovacs, *Phys. Rev.* **93** (1954), 252.
- 11) D. Ito, K. Mori and M. Yamazaki, To be published in *Nuovo Cimento*.
- 12) K. Sawada, *Prog. Theor. Phys.* **9** (1953), 455.
- 13) G. F. Chew and F. E. Low, *Phys. Rev.* **101** (1956), 1570.
- 14) D. Bodarsky, A. M. Sachs and J. Steinberger, *Phys. Rev.* **90** (1953), 997.
- 15) J. Orear, *Phys. Rev.* **98** (1955), 239.
- 16) H. A. Bethe, *Phys. Rev.* **76** (1949), 38.
- 17) Proceeding of Fifth Rochester Conference (1955).
- 18) S. Deser, W. E. Thirring and M. L. Goldberger, *Phys. Rev.* **94** (1954), 711.

*) This value may suffer some change.

Progress of Theoretical Physics, Vol. 16, No. 6, December 1956

On the Quantum Field Theory of the Interaction between the Graviton and the Matter Field, II

—Formulation without the Coordinate Condition—

Toshiei KIMURA

*Research Institute for Theoretical Physics, Hiroshima University
(Takehara-machi, Hiroshima-ken)*

(Received August 15, 1956)

The quantum theory of Einstein's linear gravitational field interacting with the elementary particles is formulated without using the coordinate condition. The canonical formalism is built up by the Dirac method of generalized Hamiltonian dynamics. It is necessary, for the requirement of being quantized in a consistent way, that the operation of creation of a spinor particle must involve the simultaneous creation of a spinor particle and of the gravitational potential around it. We show that this procedure of the quantization corresponds to the method extended from the Yang-Feldman's in the previous paper¹⁾. In the Heisenberg representation, we decompose, in the usual covariant way, the gravitational field into three parts: the transverse graviton, the gravitational potential between the matters, and the part which cannot be quantized. And also the formalism in the Heisenberg representation is transformed into the one in the interaction representation, and the same interaction Hamiltonian is obtained as in the previous paper¹⁾.

§ 1. Introduction

The previous paper¹⁾ (referred to hereafter as [I]) dealt with what role the graviton plays in the problem of the interaction between the elementary particles. There, by adopting the coordinate condition

$$\partial_\nu \gamma_{\mu\nu} \mathcal{V} = 0 \quad (1.1)$$

after Gupta²⁾, we decomposed the gravitational field into the part containing the transverse graviton (which behaves like a particle with spin 2) and the part giving the gravitational potential between the matters, and investigated their effects upon the microscopical and macroscopical phenomena. The equation (1.1) is of the same type as Lorentz condition in quantum electrodynamics. Then, when we use (1.1), the divergence difficulties of the norm of the state vector \mathcal{V} appear as in the case of quantum electrodynamics. To avoid these difficulties, we must quantize the gravitational field by using the indefinite metric in Hilbert space. But, the concept of the indefinite metric seems to us artificial and inadequate. It is the aim of the present paper to quantize, without using the coordinate condition, the approximate equation of the Einstein gravitational field interacting with the elementary particles. As the physical properties of the quantized gravitational field and its effects on the interactions between the elementary particles were

dealt with in [I], we shall not here give an account of these points.

The formulation of quantum electrodynamics without Lorentz condition has been proposed by Valatin³⁾, Dirac⁴⁾ and Ozaki⁵⁾. They started from the equation

$$\square A_\mu - \partial_\mu \partial_\lambda A_\lambda = -J_\mu$$

which is invariant under the general gauge transformations ($A_\mu \rightarrow A_\mu + \partial_\mu \cdot l$, $\square \cdot l \neq 0$). The electromagnetic potential A_μ satisfying the above equation has only three independent components, though Lorentz condition is not imposed. Therefore, taking out Coulomb potential by the suitable method and ignoring the undetermined component, we can see that the formalism without Lorentz condition is equivalent to the usual one after the longitudinal part is eliminated by the canonical transformation and Lorentz condition. (We can show that the gauge invariant Dirac operator of one electron emission¹⁾ is nothing but that obtained from the usual one by Schwinger's canonical transformation.⁶⁾)

Analogously, we can show also in the gravitational field that the formulation without using the coordinate condition gives the same description as that in [I], when we start from the gravitational equation which is invariant under the general gauge transformation. In contrast to the case of electromagnetic field, complicated features will appear, when the graviton interacts with Fermion, owing to the fact that the interaction Lagrangian containing the same rank of derivatives as that of Lagrangian of the free Fermion case. (If we use Kemmer's β -formalism for Boson, the interaction term between the graviton and Boson does not contain the derivatives of the field quantities. See [I].)

Since we have dropped the coordinate condition (1.1), no restriction is imposed on A_μ ($\mu=1, 2, 3, 4$) which appeared in the usual gauge transformations⁷⁾:

$$\gamma_{\mu\nu} \rightarrow \gamma_{\mu\nu} + \partial_\mu A_\nu + \partial_\nu A_\mu - \delta_{\mu\nu} \partial_\lambda A_\lambda. \quad (1.2)$$

In other words, our theory should be invariant under the general gauge transformations. In § 2, we obtain such a Lagrangian density which satisfies this requirement.

The coordinate condition has been needed to express the time derivatives of field quantities in terms of the independent functions of momenta. Accordingly, so long as this condition is dropped, the time derivatives of field quantities can not be expressed in terms of the independent functions of momenta and therefore the usual canonical formalism is not applicable. But, we shall show that the canonical formalism can be built up by the Dirac method⁸⁾. This is the main aim of § 3. As mentioned already, when we take Bose particle as the matter field interacting with the graviton there occurs no more complicated feature than in the case of the free graviton. So we omit the canonical formalism for the system of the graviton and Bose particle, and shall discuss the formalism for the system of the graviton and the spinor particle. It is shown that we must take a creation operator of spinor particle which accompanies its gravitational potential (not the transverse graviton!) in order that the procedure of quantization may be proceeded consistently. We can also see that this procedure corresponds to the method extended from the Yang-Feldman's⁹⁾ used in the previous paper. [I]. In quantum electrodynamics, $\text{div } E = -J_4$ can be regarded as Dirac's χ equation (the field equation which contains no

time derivative of field quantity, see § 3) and is used as the generator of the general gauge transformations. On the contrary, for the gravitational field we have no χ equation except in the case of free graviton owing to the existence of time derivatives in the source density. However, replacing the time derivative by space derivatives with the aid of field equation for the spinor particle, we are able to obtain χ equations.

In § 4, starting from the Heisenberg representation and using the familiar covariant formalism, we decompose $\gamma_{\alpha\beta}$ into the transverse part and the gravitational potential without using the coordinate condition. We obtain the same interaction Hamiltonian in the interaction representation as in [1].

§ 2. Gauge invariant Lagrangian density

We shall find the Lagrangian density which is invariant under the general gauge transformation (1.2), with the aid of the 'vier Beine' $H_{\alpha\lambda}^{(10)}$ (the index α represents the α -th leg, whereas the index λ represents the λ -th component of the vector with respect to the coordinate system in the curved space with the metric $g_{\mu\nu} = H_{\mu}^{\alpha} H_{\nu}^{\beta}$). The usual Lagrangian density for Einstein's gravitational field

$$L_G = - (1/\kappa^2) \sqrt{g} g^{\lambda\mu} (\partial_\nu \Gamma_{\lambda\mu}^\nu - \partial_\mu \Gamma_{\lambda\nu}^\nu + \Gamma_{\lambda\mu}^\rho \Gamma_{\rho\nu}^\nu - \Gamma_{\lambda\nu}^\rho \Gamma_{\rho\mu}^\nu) \quad (2.1)$$

where $\Gamma_{\lambda\mu}^\nu$ is the Christoffel three-index symbol and $\kappa^2 = 16\pi G$ (G : Newton's gravitational constant), can be expressed in terms of $H_{\alpha\lambda}$ as*

$$L_G = - (H/\kappa^2) \{ \partial_{[\lambda} H_{\alpha\rho]} \partial_{[\mu} H_{\beta\sigma]} H_{\alpha}^{\rho} H_{\beta}^{\sigma} H_{\gamma}^{\lambda} H_{\gamma}^{\mu} - \frac{1}{2} \partial_{[\lambda} H_{\alpha\sigma]} \partial_{[\mu} H_{\beta\rho]} H_{\alpha}^{\sigma} H_{\beta}^{\rho} H_{\gamma}^{\lambda} H_{\gamma}^{\mu} \\ - \frac{1}{4} \partial_{[\lambda} H_{\alpha\rho]} \partial_{[\mu} H_{\beta\nu]} H_{\gamma}^{\lambda} H_{\gamma}^{\mu} H_{\delta}^{\rho} H_{\delta}^{\nu} \} + (2/\kappa^2) \partial_\lambda \{ H_{\alpha}^{\lambda} \partial_\rho (H_{\alpha}^{\rho} H) \} \quad (2.2)$$

where $H = \sqrt{g}$. To investigate the quantum theory of the interaction between the graviton and the elementary particle, it is convenient to regard the Einstein theory as a theory of gravitation in a flat space with a Lagrangian density containing an infinite number of terms. Since (2.2) is invariant under the general Lorentz transformation (the parameters of transformations depend on the coordinates) in $\bar{\alpha}$ -system, we can assume temporarily $H_{\mu\nu}$ having the form

$$H_{\mu\nu} = \delta_{\mu\nu} + \kappa h_{\mu\nu}. \quad (2.3)$$

Expanding (2.2) in powers of κ , we obtain in the lowest order approximation

$$L_G = - \{ \partial_{[\rho} h_{\mu\mu]} \partial_{[\rho} h_{\nu\nu]} - \frac{1}{2} \partial_{[\rho} h_{\mu\nu]} \partial_{[\rho} h_{\nu\mu]} - \frac{1}{4} \partial_{[\rho} h_{\mu\nu]} \partial_{[\rho} h_{\nu\mu]} \} \quad (2.4)$$

apart from the divergence term which contributes nothing to the equation of motion for $h_{\mu\nu}$. The Lagrangian density (2.4) is evidently invariant under the gauge transformations

$$\delta h_{\mu\nu} = \partial_\nu A_\mu \quad (2.5)$$

where A_μ are arbitrary functions.

* In the following, we make use of the symbols: $A_{(\mu,\nu)} = A_{\mu,\nu} + A_{\nu,\mu}$, $A_{[\mu,\nu]} = A_{\mu,\nu} - A_{\nu,\mu}$, and take the units $\hbar = c = 1$ and space-time coordinates $x_1, x_2, x_3, x_4 = it$.

In order to proceed in parallel with the discussion in [I], we introduce $\gamma_{\mu\nu}$ by the equation

$$\gamma_{\mu\nu} = h_{\mu\nu} + h_{\nu\mu} - \partial_{\mu\nu} h_{\lambda\lambda}. \quad (2.6)$$

The $\gamma_{\mu\nu}$ is related with $g_{\mu\nu}$ by the relation $g_{\mu\nu} = \partial_{\mu\nu} + \kappa(\gamma_{\mu\nu} - \frac{1}{2}\partial_{\mu\nu}\gamma_{\lambda\lambda})$. Let $a_{\mu\nu}$ be the antisymmetric part of $h_{\mu\nu}$, i.e.,

$$a_{\mu\nu} = h_{\mu\nu} - h_{\nu\mu}. \quad (2.7)$$

In terms of $\gamma_{\mu\nu}$ and $a_{\mu\nu}$, (2.4) becomes

$$L_G = -\frac{1}{4}[\partial_\lambda\gamma_{\mu\nu} \cdot \partial_\lambda\gamma_{\mu\nu} - 2\partial_\lambda\gamma_{\mu\rho} \cdot \partial_\rho\gamma_{\mu\lambda} - \frac{1}{2}\partial_\lambda\gamma_{\rho\rho} \cdot \partial_\lambda\gamma_{\rho\rho} + \partial_\lambda(\gamma_{\rho\mu} + a_{\rho\mu})\partial_\rho(\gamma_{\lambda\mu} + a_{\lambda\mu}) \\ - \partial_\rho(\gamma_{\rho\mu} + a_{\rho\mu})\partial_\lambda(\gamma_{\lambda\mu} + a_{\lambda\mu})]. \quad (2.8)$$

In what follows, we take $\gamma_{\mu\nu}$ and $a_{\mu\nu}$ as field quantities. The above Lagrangian density is invariant under the gauge transformations:

$$\delta\gamma_{\mu\nu} = \partial_{[\mu}A_{\nu]} - \delta_{\mu\nu}\partial_\lambda A_\lambda, \quad \delta a_{\mu\nu} = \partial_{[\mu}A_{\nu]} \quad (2.9)$$

which are the consequence of (2.5).

In the presence of matter field, we add the interaction Lagrangian density L_I to L_G :

$$L_I = -(\kappa/2)(\gamma_{\mu\nu} - \frac{1}{2}\partial_{\mu\nu}\gamma_{\lambda\lambda})T_{\mu\nu} \quad (2.10)$$

where $T_{\mu\nu}$ is the symmetrical energy momentum tensor for the matter field. The field equation for $\gamma_{\mu\nu}$ is derived from the Lagrangian density (2.8) with the interaction term (2.10)

$$\square\gamma_{\mu\nu} - \partial_\mu\partial_\rho\gamma_{\rho\nu} - \partial_\nu\partial_\rho\gamma_{\rho\mu} - \frac{1}{2}\partial_{\mu\nu}\square\gamma_{\lambda\lambda} = \kappa T_{\mu\nu} - (\kappa/2)\partial_{\mu\nu}T_{\lambda\lambda}. \quad (2.11)$$

By contracting μ and ν , we have

$$\square\gamma_{\lambda\lambda} + 2\partial_\lambda\partial_\rho\gamma_{\lambda\rho} = \kappa T_{\lambda\lambda}.$$

Inserting this equation into (2.11), we obtain

$$\square\gamma_{\mu\nu} - \partial_\nu\partial_\rho\gamma_{\rho\mu} - \partial_\mu\partial_\rho\gamma_{\rho\nu} + \partial_{\mu\nu}\partial_\lambda\partial_\rho\gamma_{\lambda\rho} = \kappa T_{\mu\nu}. \quad (2.12)$$

For $a_{\mu\nu}$ we obtain only a trivial equation $0=0$. That is, $a_{\mu\nu}$ are completely undetermined.

§ 3. Canonical formalism

As we have dropped the coordinate condition, the time derivatives of the field quantities are not independent functions of momenta. To such a dynamical system the usual canonical formalism is not applicable. Therefore, we refer to the Dirac's method of generalized Hamiltonian dynamics^{9),*} We shall first summarize the Dirac's method with a few modification so as to be applicable to our problem.

* R. Utiyama once attempted to give a different method for the quantization of Einstein's linear gravitational field interacting with the vector meson field.¹⁰⁾

On the analogy of the particle dynamics, the field quantities γ_A ($A=1, 2, \dots, p$) and its space-like derivatives $\partial_\mu^s \gamma_A$ are called coordinates, and time-like derivatives $\dot{\gamma}_A = (n\partial)\gamma_A$ velocities. Momentum densities Π^A which are canonically conjugate to γ_A are introduced by the ordinary definitions from the Lagrangian density $L(\gamma_A, \partial_\mu \gamma_A)$

$$\Pi^A = \partial L / \partial \dot{\gamma}_A. \quad (\text{D} \cdot 1)$$

If Π^A are not independent functions of the velocities, the velocities can be eliminated from equations (D·1) and some equations (called ϕ equations) involving only coordinates and momenta

$$\phi_m(\Pi^A, \gamma_A, \partial_\mu^s \Pi^A, \partial_\mu^s \gamma_A) = 0, \quad m=1, 2, \dots, q \quad (\text{D} \cdot 2)$$

can be obtained.

The equations are distinguished in two kinds. One is called strong if it remains valid after the infinitesimal variations of coordinates, velocities and momenta are performed independently of each other. Weak equations are those which in general do not remain valid after such a variation. Strong equations are written with the sign \equiv , weak equations with the usual sign $=$. (D·1) and (D·2) are weak equations. If we multiply two weak equations together the resulting one turns out to be strong. For instance, from (D·1)

$$(\Pi^A - \partial L / \partial \dot{\gamma}_A) (\Pi^A - \partial L / \partial \dot{\gamma}_A) \equiv 0. \quad (\text{D} \cdot 3)$$

Hamiltonian density, defined in the usual manner :

$$H \equiv \Pi^A \dot{\gamma}_A - L \quad (\text{D} \cdot 4)$$

can be expressed in terms of ϕ_m as

$$H \equiv \mathfrak{H}(\Pi^A, \gamma_A, \partial_\mu^s \gamma_A) + v_m \phi_m \quad (\text{D} \cdot 5)$$

where v_m are functions of the coordinates and velocities. By taking the time t to be an extra coordinate γ_0 and using the equation $\dot{\gamma}_0 = 1$, the Lagrangian density can be regarded to be homogeneous of the first degree in all velocities, including $\dot{\gamma}_0$. In this case, H can be expressed as

$$H \equiv v_m \phi_m. \quad (\text{D} \cdot 6)$$

Among the field equations, the one which can be expressed in terms of the momenta and the coordinates is called χ equation :

$$\chi_t = 0. \quad (\text{D} \cdot 7)$$

A ϕ_m is defined to be a first class ϕ if its Poisson bracket (P.b.) with every ϕ and χ vanishes, and a ϕ_m which does not satisfy this condition is called a second class ϕ . As the equation of motion is not consistent with the equation $\phi=0$ where ϕ is the second class, the Hamiltonian form must be modified. For this purpose, the following P.b. is introduced :

$$(\xi, \eta')^* \equiv (\xi, \eta') + \iint d\mathbf{x}'' d\mathbf{x}''' (\xi, \theta_s'') C_{ss'}(\theta_s''', \eta') \quad (\text{D} \cdot 8)$$

where θ_s are the second class ϕ , (\tilde{z}, γ') the usual P.b. and

$$C_{ss'} \equiv \text{co-factor of } (\theta_s'', \theta_{s'}''')/D \quad (\text{D} \cdot 9)$$

with

$$D = \begin{vmatrix} 0, & (\theta_1'', \theta_2'''), & (\theta_1'', \theta_3'''), \dots, & (\theta_1'', \theta_s''') \\ (\theta_2'', \theta_1'''), & 0, & (\theta_2'', \theta_3'''), \dots, & (\theta_2'', \theta_s''') \\ \vdots & \vdots & \vdots & \vdots \\ (\theta_s'', \theta_1'''), & & & 0 \end{vmatrix}.$$

By using the above Dirac P.b. instead of the usual one, each of the second class ϕ 's can be regarded to vanish in the strong sense, and the degree of freedom can be reduced.

The quantization is carried out on the following principles.

i) New P.b. relations between the classical variables, $(\gamma_A(\mathbf{x}), H^{A'}(\mathbf{x}'))^*$, correspond to commutation or anticommutation relations between the operators: $-i[\gamma_A(\mathbf{x}), H^{A'}(\mathbf{x}')]_{\mp} \equiv -i(\gamma_A(\mathbf{x}) H^{A'}(\mathbf{x}') \mp H^{A'}(\mathbf{x}') \gamma_A(\mathbf{x}))$.

ii) The first class ϕ and χ equations between the classical variables, $\phi(H^A, \gamma_A) = 0$ and $\chi(H^A, \gamma_A) = 0$, correspond to linear conditions on the state vector ψ : $\phi(H^A, \gamma_A)\psi = 0$, $\chi(H^A, \gamma_A)\psi = 0$.

We shall now apply the above Dirac method to the case of free graviton. The canonical momenta $H_{\mu\nu}$ and $A_{\mu\nu}$ conjugate to $\gamma_{\mu\nu}$ and $a_{\mu\nu}$ are derived from (2.8), respectively

$$H_{\mu\nu} = -n_\lambda \partial L_G / \partial (\partial_\lambda \gamma_{\mu\nu}) \equiv \frac{1}{2} [-n_\mu n_\nu \partial \gamma_{nn} + \partial \gamma_{\mu\nu}^s - \frac{1}{2} \partial_\mu^s (\gamma_{\nu n} + a_{\nu n}) - \frac{1}{2} n_{(\mu} \partial_\rho^s (\gamma_{\rho\nu)} + a_{\rho\nu}) - \frac{1}{2} \partial_{\mu\nu} \partial \gamma_{\lambda\lambda}], \quad (3.1)$$

$$A_{\mu\nu} = \frac{1}{4} [\partial_{[\mu}^s (\gamma_{n\nu]}^s + a_{n\nu}^s) - n_{[\mu} \partial_\rho^s (\gamma_{\rho\nu]}^s + a_{\rho\nu}^s) - \partial_{[\mu}^s n_{\nu]} \gamma_{nn}] \quad (3.2)$$

where

$$\begin{cases} \gamma_{\mu n} \equiv n_\nu \gamma_{\mu\nu}, & \gamma_{nn} \equiv n_\mu n_\nu \gamma_{\mu\nu}, & \gamma_{\mu n}^s \equiv \gamma_{\mu n} + n_\mu \gamma_{nn}, \\ \gamma_{\mu\nu}^s \equiv \gamma_{\mu\nu} + n_{(\mu} \gamma_{n\nu)} + n_\mu n_\nu \gamma_{nn} \equiv \gamma_{\mu\nu} + n_{(\mu} \gamma_{n\nu)}^s - n_\mu n_\nu \gamma_{nn}, \\ \partial \equiv (n\partial), & \partial_\mu^s \equiv \partial_\mu + n_\mu \partial. \end{cases} \quad (3.3)$$

Equation (3.2) is clearly ϕ equation because of involving no time derivative. We denote it by

$$\phi_{\mu\nu} \equiv A_{\mu\nu} - \frac{1}{4} [\partial_{[\mu}^s (\gamma_{n\nu]}^s + a_{n\nu}^s) - n_{[\mu} \partial_\rho^s (\gamma_{\rho\nu]}^s + a_{\rho\nu}^s) - \partial_{[\mu}^s n_{\nu]} \gamma_{nn}] = 0. \quad (3.4)$$

From (3.1) and (3.2),

$$H_{\mu n}^s = -A_{\mu n}^s = \frac{1}{4} [-\partial_\mu^s \gamma_{nn} + \partial_\rho^s (\gamma_{\rho\mu}^s + a_{\rho\mu}^s)]. \quad (3.5)$$

Then we obtain ϕ equations

$$\phi_\mu^s \equiv H_{\mu n}^s + A_{\mu n}^s = 0. \quad (3.6)$$

Similary

$$\phi_n \equiv H_{\lambda\lambda}^s + H_{nn} = 0. \quad (3.7)$$

We can show that χ equations exist. From the equation (2.12), we obtain

$$4\gamma_{\mu n} - \partial_\lambda^s \partial \gamma_{\mu\lambda}^s + \partial_\mu^s \partial \gamma_{nn}^s - \partial_\mu^s \partial_\lambda^s \gamma_{\lambda n}^s + n_\mu \partial_\lambda^s \partial_\rho^s \gamma_{\lambda\rho}^s = 0 \quad (3.8)$$

where $J \equiv \partial_\mu^s \partial_\mu^s \equiv \square + \partial^2$. Expressing $\partial \gamma_{\mu\lambda}^s$ and $\partial \gamma_{nn}^s$ in terms of momenta with the aid of (3.1), we obtain χ equation

$$\chi_\mu^s \equiv \partial_\nu^s (H_{\mu\nu}^s + A_{\mu\nu}^s) + \partial_\mu^s H_{nn}^s = 0. \quad (3.9)$$

Similarly

$$\chi_n \equiv 4\gamma_{nn} - \partial_\lambda^s \partial_\rho^s \gamma_{\lambda\rho}^s = 0. \quad (3.10)$$

Thus, requisite ϕ and χ equations are obtained.

We shall next transform the Hamiltonian

$$H \equiv \partial \gamma_{nn}^s H_{nn}^s - 2\partial \gamma_{n\mu}^s H_{n\mu}^s + \partial \gamma_{\mu\nu}^s H_{\mu\nu}^s - 2\partial a_{n\mu}^s A_{n\mu}^s + \partial a_{\mu\nu}^s A_{\mu\nu}^s - L_G \quad (3.11)$$

into the form of (D.5). For this transformation, we must use the strong equation alone, such as (D.3). Adding the strong equation:

$$[H_{\mu\nu}^s + A_{\mu\nu}^s - \frac{1}{2}(\partial \gamma_{\mu\nu}^s - \partial_\lambda^s \gamma_{\mu\lambda}^s - \partial_\nu^s a_{\mu\lambda}^s - \frac{1}{2}\partial_\mu^s \partial_\nu^s \gamma_{\lambda\lambda}^s)]^2 + [H_{nn}^s + \frac{1}{4}\{3\gamma_{nn}^s - \partial \gamma_{\lambda\lambda}^s - 2\partial_\rho^s (\gamma_{\rho n}^s + a_{\rho n}^s)\}]^2 \\ - \frac{1}{2}[H_{\lambda\lambda}^s - H_{nn}^s - \frac{1}{2}\{3\gamma_{nn}^s - \partial \gamma_{\lambda\lambda}^s - 2\partial_\rho^s (\gamma_{\rho n}^s + a_{\rho n}^s)\}]^2 \equiv 0 \quad (3.12)$$

to (3.11), we obtain after straightforward calculations

$$H \equiv \mathfrak{H} - \partial_\mu \gamma_{n\mu}^s \cdot \phi_n - 2\partial \gamma_{n\lambda}^s \cdot \phi_\lambda^s + \partial a_{\mu\nu}^s \cdot \phi_{\mu\nu} \quad (3.13)$$

where

$$\mathfrak{H} \equiv (H_{\mu\nu}^s + A_{\mu\nu}^s + \delta_{\mu\nu}^s H_{nn}^s) \partial_\nu^s (\gamma_{\mu n}^s + a_{\mu n}^s) + H_{\mu\nu}^s H_{\mu\nu}^s - H_{\lambda\lambda}^s H_{\lambda\lambda}^s + A_{\mu\nu}^s A_{\mu\nu}^s - L_G^s \\ + \frac{1}{4}(\frac{1}{2}\partial_\lambda^s \gamma_{n\mu}^s \cdot \partial_\lambda^s \gamma_{nn}^s - 2\partial_\lambda^s \gamma_{n\mu}^s \cdot \partial_\lambda^s \gamma_{\mu\lambda}^s + 2\partial_\lambda^s \gamma_{n\mu}^s \cdot \partial_\mu^s \gamma_{n\lambda}^s + \partial_\lambda^s \gamma_{nn}^s \cdot \partial_\lambda^s \gamma_{\rho\rho}^s) \quad (3.14)$$

in which L_G^s means the term of L_G having only the space-like indices, i.e.,

$$L_G^s \equiv -\frac{1}{4}[\partial_\lambda^s \gamma_{\mu\nu}^s \cdot \partial_\lambda^s \gamma_{\mu\nu}^s - 2\partial_\lambda^s \gamma_{\mu\rho}^s \cdot \partial_\rho^s \gamma_{\mu\lambda}^s - \frac{1}{2}\partial_\lambda^s \gamma_{\rho\rho}^s \cdot \partial_\lambda^s \gamma_{\rho\rho}^s \\ + \partial_\lambda^s (\gamma_{\rho\mu}^s + a_{\rho\mu}^s) \partial_\rho^s (\gamma_{\lambda\mu}^s + a_{\lambda\mu}^s) - \partial_\rho^s (\gamma_{\rho\mu}^s + a_{\rho\mu}^s) \partial_\lambda^s (\gamma_{\lambda\mu}^s + a_{\lambda\mu}^s)]. \quad (3.15)$$

In order to study whether the above ϕ 's belong to the first or to the second class, we define P.b.'s at $t=t'$

$$(H_{\mu\nu}, \gamma_{\lambda\rho}'^s) \equiv -\frac{1}{2}\partial_{[\mu\lambda}\partial_{\nu]\rho} \partial(x-x'), \\ (A_{\mu\nu}, a_{\lambda\rho}') \equiv -\frac{1}{2}\partial_{[\mu\lambda}\partial_{\nu]\rho} \partial(x-x'). \quad (3.16)$$

As is easily seen, each ϕ has zero P.b. with every ϕ and χ . That is, all ϕ and χ belong to the first class.* Therefore, the quantization is easily carried out by putting ϕ and

* χ_μ^s defined by (3.9) and ϕ_n by (3.7) play the role of the generator of gauge transformation. In fact, if we take the generator as

$$G = - \int (d^3x) \{ 2A_r \{ \partial_s (H_{rs} + A_{rs}) + \partial_r H_{00} \} + \partial_\mu A_\mu \{ H_{rr} + H_{00} \} \}$$

in the special reference system $n_\mu = (0, 0, 0, i)$, $\delta \gamma_{ij} = (\gamma_{ij}, G)$ and $\delta a_{ij} = (a_{ij}, G)$ are nothing but the gauge transformations (2.9).

χ equations as the linear conditions on the state vector Ψ :

$$[H_{\lambda\lambda}^s + H_{nn}] \Psi = 0, \quad (3 \cdot 17_1)$$

$$[H_{\mu\mu}^s + A_{\mu n}^s] \Psi = 0, \quad (3 \cdot 17_2)$$

$$[A_{\mu\nu} - \frac{1}{4} \{ \partial_{[\mu}^s (\gamma_{\nu]}^s + a_{\nu]}^s) - n_{[\mu} \partial_{\rho}^s (\gamma_{\rho\nu]}^s + a_{\rho\nu]}^s) - \partial_{[\mu}^s n_{\nu]} \gamma_{nn}^s \}] \Psi = 0, \quad (3 \cdot 17_3)$$

$$[\partial_{\nu}^s (H_{\mu\nu}^s + A_{\mu\nu}^s) + \partial_{\mu}^s H_{nn}] \Psi = 0, \quad (3 \cdot 17_4)$$

$$[A_{\gamma_{nn}} - \partial_{\lambda}^s \partial_{\rho}^s \gamma_{\lambda\rho}^s] \Psi = 0. \quad (3 \cdot 17_5)$$

We shall here introduce a new field quantity, γ , because of the following two reasons :

i) Among the ϕ equations, only ϕ_n defined by (3·7) is the relation among $H_{\mu\mu}$ themselves, while the others contain $A_{\mu\nu}$. So we should like to change the weak equation (3·7) into the strong one. ii) It is desirable to develop the theory in parallel with [I]. The new γ is defined through $\chi^{(2)}$ equation

$$\chi^{(2)} \equiv \gamma - \gamma_{\lambda\lambda} = 0. \quad (3 \cdot 18)$$

And further, we assume that the canonical momentum density conjugate to γ is

$$A = -\frac{1}{4} \partial \gamma_{\lambda\lambda}. \quad (3 \cdot 19)$$

Since the P.b. of ϕ_n with $\chi^{(2)}$ is not zero, the ϕ_n and $\chi^{(2)}$ belong to the second class. Therefore, we shall use the Dirac P.b.'s. By taking $\theta_1 = \phi_n$ and $\theta_2 = \chi^{(2)}$ for θ 's in (D·8), $C_{ss'}$ defined by (D·9) are expressed as

$$(C_{ss'}) \equiv \begin{pmatrix} 0 & \frac{1}{2} \partial (\mathbf{x}'' - \mathbf{x}''') \\ -\frac{1}{2} \partial (\mathbf{x}'' - \mathbf{x}''') & 0 \end{pmatrix}.$$

Hence, the new P.b.'s defined by (D·8) become

$$(H_{\mu\nu}, \gamma'_{\lambda\rho})^* \equiv -\frac{1}{2} [\partial_{(\mu\lambda} \partial_{\nu)\rho} - \delta_{\mu\nu} (\partial_{\lambda\rho} + 2n_{\lambda} n_{\rho})] \partial (\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 20_1)$$

$$(A_{\mu\nu}, a'_{\lambda\rho})^* \equiv -\frac{1}{2} \partial_{[\mu\lambda} \partial_{\nu]\rho} \partial (\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 20_2)$$

$$(A, \gamma'_{\lambda\rho})^* \equiv -\frac{1}{2} [\partial_{\lambda\rho} + 2n_{\lambda} n_{\rho}] \partial (\mathbf{x} - \mathbf{x}'), \quad (A, \gamma')^* \equiv -\partial (\mathbf{x} - \mathbf{x}'). \quad (3 \cdot 20_3)$$

The other new P.b.'s not presented here are zero. Thus, if we work only with the new P.b.'s, we may assume that ϕ_n and $\chi^{(2)}$ vanish in the strong sense without getting a contradiction, because new P.b.'s of ϕ_n and $\chi^{(2)}$ with any dynamical variable are zero. Accordingly, we can reduce the freedom of H_{nn} , γ_{nn} and supplementary condition (3·17₁). Putting

$$\widetilde{H}_{\mu\nu} \equiv H_{\mu\nu} + \delta_{\mu\nu} A, \quad (3 \cdot 21)$$

$$\widetilde{A} \equiv A - H_{nn} \equiv -\widetilde{H}_{nn}, \quad (3 \cdot 22)$$

we have

$$(\widetilde{H}_{\mu\nu}, \gamma'_{\lambda\rho})^* \equiv -\frac{1}{2} \partial_{(\mu\lambda}^s \partial_{\nu)\rho}^s \partial (\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 23)$$

$$(\widetilde{A}, \gamma'_{\lambda\rho})^* \equiv 0, \quad (\widetilde{A}, \gamma')^* \equiv -\partial (\mathbf{x} - \mathbf{x}'). \quad (3 \cdot 24)$$

These new P.b.'s have the familiar forms.

Hitherto we have discussed only the case of free graviton. We shall next consider the case where the matter is present. When we take Bose particle as a matter field, there occurs no more complicated feature than in the case of free graviton and the quantization procedure is followed in the way similar to that in the free case. This corresponds to the following fact. For the system of Boson and graviton, there exist unitary transformations which connect the canonical variables in the interacting system with those in the free system, while such transformations do not exist for the system of Fermion and graviton. Then, we shall deal in detail with the case of the graviton interacting with the spinor particle only.

As the Lagrangian density we take

$$L \equiv L_G - \frac{1}{2} (\bar{\psi} \alpha_\mu \partial_\mu \psi - \partial_\mu \bar{\psi} \alpha_\mu \psi + 2m \bar{\psi} \psi) - (\kappa/4) h_{\mu\nu} (\bar{\psi} \alpha_\nu \partial_\mu \psi - \partial_\mu \bar{\psi} \alpha_\nu \psi) \quad (3 \cdot 25)$$

where $h_{\mu\nu} \equiv \gamma_{\mu\nu} - \frac{1}{2} \delta_{\mu\nu} \gamma$ and $\bar{\psi} \equiv \psi^\dagger \alpha_4$ (ψ^\dagger : Hermitian conjugate of ψ). The momentum densities, Π and Π^+ , canonically conjugate to ψ and ψ^+ are

$$\Pi = (i/2) \psi^\dagger [1 - (\kappa/2) (\alpha_n \alpha_\lambda \gamma_{n\lambda} + \frac{1}{2} \gamma)], \quad (3 \cdot 26_1)$$

$$\Pi^+ = - (i/2) [1 - (\kappa/2) (\alpha_n \alpha_\lambda \gamma_{n\lambda} + \frac{1}{2} \gamma)] \psi. \quad (3 \cdot 26_2)$$

As the equations (3·26) do not contain time derivatives, they are also ϕ equations. We write them as

$$\phi_\alpha^{(2)} \equiv \Pi_\alpha - \frac{i}{2} \left[\psi_\alpha^\dagger \left(1 - \frac{\kappa}{4} \gamma \right) - (\kappa/2) (\psi^\dagger \alpha_n \alpha_\lambda) \gamma_{n\lambda} \right] = 0, \quad (3 \cdot 27_1)$$

$$\phi_\beta^{+(2)} \equiv \Pi_\beta^+ + \frac{i}{2} \left[\left(1 - \frac{\kappa}{4} \gamma \right) \psi_\beta - (\kappa/2) \gamma_{n\lambda} (\alpha_n \alpha_\lambda \psi)_\beta \right] = 0. \quad (3 \cdot 27_2)$$

The ϕ_μ^s and $\phi_{\mu\nu}$ have the same forms as those in the free graviton case. Setting up the P.b.'s

$$(\Pi_\alpha, \psi'_\beta) \equiv (\Pi_\alpha^+, \psi_\beta^{+'}) \equiv -\delta_{\alpha\beta} \delta(\mathbf{x} - \mathbf{x}') \quad (3 \cdot 28)$$

we obtain

$$(\phi_\alpha^{+(2)}(\mathbf{x}), \phi_\beta^{(2)}(\mathbf{x}')) \equiv -i E_{\alpha\beta} \delta(\mathbf{x} - \mathbf{x}') \quad (3 \cdot 29)$$

where

$$E \equiv 1 - (\kappa/2) \{ \alpha_n \alpha_\lambda \gamma_{n\lambda} + \frac{1}{2} \gamma \}. \quad (3 \cdot 30)$$

The $\phi_\alpha^{(2)}$ and $\phi_\beta^{(2)}$ are then the second class ϕ 's. And further they have non-vanishing P.b.'s with ϕ_μ^s given by (3·6). Therefore, it is necessary to turn $\phi_\alpha^{(2)} = 0$ and $\phi_\beta^{+(2)} = 0$ to the strong equations. For this purpose, we shall adopt the Dirac P.b. (D·8).

Taking

$$\theta_\alpha \equiv \phi_\alpha^{+(2)}, \quad \theta_{4+\beta} \equiv \phi_\beta^{(2)} \quad (3 \cdot 31)$$

as θ_s , we obtain

$$(C_{ss'}) \equiv \begin{pmatrix} 0 & i(1/E) \alpha_{\beta} \partial(\mathbf{x}'' - \mathbf{x}'''), \\ -i(1/E) \alpha_{\beta} \partial(\mathbf{x}'' - \mathbf{x}'''), & 0 \end{pmatrix}. \quad (3 \cdot 32)$$

New P.b.'s are then

$$(\psi_{\alpha}, \psi_{\beta}^{+'})^{*} \equiv -i(1/E) \alpha_{\beta} \partial(\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 33_1)$$

$$(H_{n\mu}^s, H_{n\nu}^{s'})^{*} \equiv i(\kappa^2/64) \psi^{+} \alpha_n \alpha_{\mu} (1/E) \alpha_n \alpha_{\nu} \psi \partial(\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 33_2)$$

$$(H_{n\mu}^s, \psi')^{*} \equiv (\kappa/8) (1/E) \alpha_n \alpha_{\mu} \psi \partial(\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 33_3)$$

$$(H_{n\mu}^s, \psi^{+'})^{*} \equiv (\kappa/8) \psi^{+} \alpha_n \alpha_{\mu} (1/E) \partial(\mathbf{x} - \mathbf{x}'), \quad (3 \cdot 33_4)$$

$$(A, \psi')^{*} \equiv -(\kappa/8) (1/E) \psi \partial(\mathbf{x} - \mathbf{x}'),$$

$$(A, \psi^{+'})^{*} \equiv -(\kappa/8) \psi^{+} (1/E) \partial(\mathbf{x} - \mathbf{x}'). \quad (3 \cdot 33_5)$$

The other new P.b.'s not presented here have the same forms as the old ones. The above new P.b.'s have forms different from the usual one, because each right-hand side of (3.33) involves $\gamma_{n\lambda}$ and γ through E and further $H_{n\mu}^s$ have not zero P.b.'s with ψ and $H_{n\nu}^{s'}$. Therefore we must extend the Dirac method so as to give the familiar P.b.'s. For this purpose, we shall use the following quantities instead of the original one:

$$\underline{\psi} \equiv \sqrt{E} \psi, \quad \underline{\psi}^{+} \equiv \psi^{+} \sqrt{E} \quad (3 \cdot 34)$$

with

$$\sqrt{E} \equiv 1 - (\kappa/4) (\alpha_n \alpha_{\lambda} \gamma_{n\lambda} + \frac{1}{2} \gamma) - (\kappa^2/32) (\alpha_n \alpha_{\lambda} \gamma_{n\lambda} + \frac{1}{2} \gamma)^2 + O(\kappa^3), \quad (3 \cdot 35)$$

and

$$\underline{H}_{n\mu}^s \equiv H_{n\mu}^s - (\kappa^2/64) \underline{\psi}^{+} \{ \alpha_n \alpha_{\lambda} \gamma_{n\lambda} \cdot \alpha_n \alpha_{\mu} - \alpha_n \alpha_{\mu} \cdot \alpha_n \alpha_{\lambda} \gamma_{n\lambda} \} \underline{\psi} + O(\kappa^3). \quad (3 \cdot 36)$$

For the new field quantities $\underline{\psi}$, $\underline{\psi}^{+}$ and $\underline{H}_{n\mu}^s$ we have the familiar P.b.'s

$$(\underline{\psi}_{\alpha}, \underline{\psi}_{\beta}^{+'})^{*} \equiv -i \partial_{\alpha\beta} \partial(\mathbf{x} - \mathbf{x}'),$$

$$(H_{n\mu}^s, H_{n\nu}^{s'})^{*} = (H_{n\mu}^s, \psi')^{*} = (H_{n\mu}^s, \psi^{+'})^{*} = (A, \psi^{+'})^{*} = (A, \psi')^{*} = 0. \quad (3 \cdot 37)$$

Thus, the quantization is carried out by replacing the Dirac P.b.'s in (3.37) by the anticommutator or commutator. But, there appears an undesirable feature. Owing to the existence of non-vanishing P.b.'s among $H_{n\mu}^s$ (in κ^2 order term), the supplementary conditions (3.17₂) become inconsistent with each other. However, since we confine here ourselves to the first order term in κ , we do not take this fact into account.

We shall consider the physical meaning of (3.37). This means that we must take, as the creation or annihilation operator of the spinor particle, the one which involves $\gamma_{n\lambda}$ and γ through \sqrt{E} given by (3.35). As will be shown in the following section (see also [I]), $\gamma_{n\lambda}$ and γ give the gravitational potential. That is, in order that one may quantize the system of the gravitational field and the spinor particle consistently, it is necessary to take the creation operator of spinor particle accompanying the gravitational potential.

At first sight our procedure seems to be different from the usual formalism, but essentially this is not the case. In order that the usual quantization method of Heisenberg-Pauli may be applied to the spinor field with the interaction term containing its derivatives, we must choose the canonical variables φ and Π so that $\partial_i \varphi$ and $\partial_i \Pi$ are not contained in the interaction Hamiltonian, and that $i\Pi = -\varphi^+$. So far as we take the original ψ as φ , we cannot eliminate time derivatives, while we can do so by using our $\underline{\psi}$ and $\underline{\psi}^+$ given in (3.36). Also, according to the method extended from the Yang-Feldman's in [I], Heisenberg operator ϕ is expressed in terms of $\psi(x/\sigma)$

$$\phi(x) = \psi(x/\sigma) + (\kappa/4) (\alpha_n \alpha_\lambda \gamma_{n\lambda} + \frac{1}{2} \gamma) \psi + (\kappa^2/32) (\alpha_n \alpha_\lambda \gamma_{n\lambda} + \frac{1}{2} \gamma)^2 \psi(x/\sigma) + O(\kappa^3), \quad (3.38)$$

and $\psi(x/\sigma)$ is used as the canonical variable. The $\psi(x/\sigma)$ is nothing but our $\underline{\psi}$.

Finally, we shall mention about χ equations in the presence of spinor particle. Equations (3.9) and (3.10) are replaced by

$$\chi_\mu^s \equiv \partial_\nu^s (H_{\mu\nu}^s + A_{\mu\nu}^s) + \partial_\nu^s H_{nn} - (\kappa/8) \{ (\bar{\psi} \alpha_n \partial_\nu^s \psi - \partial_\nu^s \bar{\psi} \alpha_n \psi) + (\bar{\psi} \alpha_n^s \partial \psi - \partial \bar{\psi} \alpha_n^s \psi) \} = 0, \quad (3.39)$$

$$\chi_n \equiv \Delta \gamma_{nn} - \partial_\lambda^s \partial_\nu^s \gamma_{\lambda\nu}^s - (\kappa/4) (\bar{\psi} \alpha_n \partial \psi - \partial \bar{\psi} \alpha_n \psi) = 0. \quad (3.40)$$

Since the equations (3.39) and (3.40) involve time derivatives, they are not χ equations. But, using the field equation for spinor particle, we express time derivatives in terms of space derivatives:

$$\begin{aligned} \bar{\psi} \alpha_\mu^s \partial \psi - \partial \bar{\psi} \alpha_\mu^s \psi &= \bar{\psi} \alpha_n \partial_\mu^s \psi - \partial_\mu^s \bar{\psi} \alpha_n \psi, \\ \bar{\psi} \alpha_n \partial \psi - \partial \bar{\psi} \alpha_n \psi &= 2m \bar{\psi} \psi \end{aligned} \quad (3.41)$$

where the first order term in κ is neglected. Therefore we can regard (3.39) and (3.40) as χ equations in the approximation up to the first order term in κ , and regard them as the generators of the general gauge transformations.

§ 4. Covariant decomposition of field quantity

Now in this section, using the usual covariant formalism, we shall develop the theory in the Heisenberg representation. The field equation for $\gamma_{\mu\nu}^*$ is given by (2.12):

$$\square \gamma_{\mu\nu} - \partial_\mu \partial_\rho \gamma_{\rho\nu} - \partial_\nu \partial_\rho \gamma_{\rho\mu} + \partial_{\mu\nu} \partial_\lambda \partial_\rho \gamma_{\lambda\rho} = \kappa T_{\mu\nu}. \quad (4.1)$$

In order to separate the transverse part from $\gamma_{\mu\nu}$, we introduce the projection operator $K_{\mu\nu}^{**}$:

$$K_{\mu\nu} \equiv \delta_{\mu\nu}^s - \partial_\mu^s \partial_\nu^s \Delta^{-1} \equiv \delta_{\mu\nu} + (n_\mu \square - \partial_\mu \partial) n_\nu \Delta^{-1} - \partial_\mu^s \partial \Delta^{-1} \quad (4.2)$$

* In this section, the bold-faced type letter stands for the operator in the Heisenberg representation.

** In quantum electrodynamics, Ozaki has used $K_{\mu\nu}' = K_{\mu\nu} + \alpha \partial(\Delta) \{n_\mu n_\nu \Delta - \partial_\mu^s \partial_\nu^s\}^{(5)}$. Since the additive term does not affect the physical effect, we drop this term here.

which satisfies the following equations :

$$\partial_\nu K_{\mu\nu} \equiv n_\nu K_{\mu\nu} \equiv 0, \quad K_{\mu\lambda} K_{\lambda\nu} \equiv K_{\mu\nu}. \quad (4.3)$$

The transverse part $\bar{\gamma}_{\mu\nu}$ is defined with the aid of $K_{\mu\nu}$:

$$\bar{\gamma}_{\mu\nu} \equiv K_{\mu\lambda} K_{\nu\rho} \bar{\gamma}_{\lambda\rho}. \quad (4.4)$$

Then the gravitational field can be decomposed into a form

$$\gamma_{\mu\nu} \equiv \bar{\gamma}_{\mu\nu} + \mathcal{E}_{\mu\nu}. \quad (4.5)$$

Multiplying $K_{\mu'\mu}$, $K_{\nu'\nu}$ from the left of (4.1) and summing over the indices μ' and ν' respectively, we obtain

$$\square \bar{\gamma}_{\mu\nu} + K_{\mu\nu} \partial_\lambda \partial_\rho \mathcal{E}_{\lambda\rho} = \kappa T_{\mu\nu} + \kappa (n_{(\mu} \square - \partial_{(\mu} \partial) J^{-1} T_{\nu)}) + \kappa (n_\mu \square - \partial_\mu \partial) (n_\nu \square - \partial_\nu \partial) J^{-2} T_{nn} \quad (4.6)$$

by putting $\mu' = \mu$ and $\nu' = \nu$. From (4.1), (4.5) and (4.6), the equation for $\mathcal{E}_{\mu\nu}$ becomes

$$\begin{aligned} \square \mathcal{E}_{\mu\nu} - \partial_\mu \partial_\lambda \mathcal{E}_{\lambda\nu} - \partial_\nu \partial_\lambda \mathcal{E}_{\mu\lambda} + (\partial_{\mu\nu} - K_{\mu\nu}) \partial_\lambda \partial_\rho \mathcal{E}_{\lambda\rho} \\ = -\kappa (n_{(\mu} \square - \partial_{(\mu} \partial) [J^{-1} T_{\nu)} + \frac{1}{2} (n_\nu) \square - \partial_\nu \partial) J^{-2} T_{nn}]. \end{aligned} \quad (4.7)$$

The solution of (4.7) can be expressed in the form :

$$\begin{aligned} \mathcal{E}_{\mu\nu} &= \mathcal{E}_{\mu\nu}(x, \sigma) - \kappa n_{(\mu} J^{-1} T_{\nu)} - \kappa n_\mu n_\nu J^{-1} T_{nn} + (\kappa/2) n_{(\mu} J^{-2} \partial_{\nu)}^s \partial T_{nn} \\ &= \mathcal{E}_{\mu\nu}(x, \sigma) + \kappa n_\mu n_\nu J^{-1} T_{nn} - \kappa n_{(\mu} J^{-1} T_{\nu)}^s + (\kappa/2) n_{(\mu} J^{-2} \partial_{\nu)}^s \partial_\lambda^s T_{n\lambda}^s \end{aligned} \quad (4.8)$$

where $\mathcal{E}_{\mu\nu}(x, \sigma)$ satisfies the equation

$$\square \mathcal{E}_{\mu\nu}(x, \sigma) - \partial_\mu \partial_\lambda \mathcal{E}_{\lambda\nu}(x, \sigma) - \partial_\nu \partial_\lambda \mathcal{E}_{\mu\lambda}(x, \sigma) + (\partial_{\mu\nu} - K_{\mu\nu}) \partial_\lambda \partial_\rho \mathcal{E}_{\lambda\rho}(x, \sigma) = 0. \quad (4.9)$$

The terms in the right-hand side of (4.8) containing the energy momentum tensor of matter give rise to the gravitational potential between the matters.

We shall next find the solution of (4.6) in the integral form. Taking into account the relation $J^{-1} D = \partial^{-2} D$, we obtain the commutation relations for incoming field $\bar{\gamma}_{\mu\nu}^{\partial^2 n}$ from the result of [I]

$$[\bar{\gamma}_{\mu\nu}^{\partial^2 n}, \bar{\gamma}_{\lambda\rho}^{\partial^2 n'}] = i d_{\mu\nu}^{\lambda\rho} D(x - x') \quad (4.10)$$

where

$$d_{\mu\nu}^{\lambda\rho} \equiv \partial_{(\mu\lambda}^s \partial_{\nu\rho)}^s - \partial_{(\mu\lambda}^s \partial_{\nu)}^s \partial_\rho^s J^{-1} - \partial_{(\mu\rho}^s \partial_{\nu)}^s \partial_\lambda^s J^{-1} + 2 \partial_\mu^s \partial_\nu^s \partial_\lambda^s \partial_\rho^s J^{-2}. \quad (4.11)$$

The appropriateness of (4.10) is also confirmed directly from (3.23). Using $d_{\mu\nu}^{\lambda\rho}$, we obtain

$$\begin{aligned} \bar{\gamma}_{\mu\nu} = \bar{\gamma}_{\mu\nu}(x, \sigma) + \int (dx') (\partial_{\mu\nu}^s - \partial_\mu^s \partial_\nu^s J^{-1}) D^s(x, x') \partial_\lambda \partial_\rho \mathcal{E}_{\lambda\rho}(x, \sigma) \\ - (\kappa/2) \int (dx') d_{\mu\nu}^{\lambda\rho} D^s(x, x') T_{\lambda\rho}(x') \end{aligned} \quad (4.12)$$

where $\bar{\gamma}_{\mu\nu}(x, \sigma)$ satisfies

$$\square \bar{\gamma}_{\mu\nu}(x, \sigma) = 0 \quad (4.13)$$

and $D^\sigma(x, x') \equiv \frac{1}{2} \{ \epsilon(\sigma, x') - \epsilon(x, x') \} D(x - x')$.

From the definition of $\bar{\gamma}_{\mu\nu}$ and the property of $K_{\mu\nu}$, $\bar{\gamma}_{\mu\nu}$ has two independent components. After transforming (4.1) into the momentum space, we regard them as the simultaneous algebraic equation which determine ten $\bar{\gamma}_{\mu\nu}$. We can show that the rank of the matrix constructed from the coefficients of $\bar{\gamma}_{\mu\nu}$ is six. Therefore only six components of $\bar{\gamma}_{\mu\nu}$ can be determined, while the remaining four are undetermined. The solution (4.8) shows that $\mathcal{E}_{\mu\nu}$ contains four components of energy momentum density $T_{n\mu}$ for the matter field. Since $T_{n\mu}$ is subject to the conservation law, only three components of $T_{n\mu}$ are independent. On the other hand, $\bar{\gamma}_{\mu\nu}$ has two independent components and γ_{nn} is replaced by the new independent field quantity, γ , introduced in § 3. Therefore the four undetermined components must be contained in $\mathcal{E}_{\mu\nu}(x, \sigma)$. The solution for $\mathcal{E}_{\mu\nu}(x, \sigma)$ can be expressed in terms of A_μ ($\mu=1, 2, 3, 4$)

$$\mathcal{E}_{\mu\nu}(x, \sigma) = \partial_{(\mu} A_{\nu)} - (\partial_{\mu\nu} - K_{\mu\nu}) \partial_\lambda A_\lambda \quad (4.14)$$

where A_μ are quite arbitrary functions. It seems, therefore, that they have some connection with A_μ in the general gauge transformations (2.9), and play the similar role as $a_{\mu\nu}$ introduced in § 2.

Finally, we shall transform the above formalism in the Heisenberg representation into the one in the interaction representation. This can be done by the same method that used in [I]. For instance, taking the scalar meson field V as a matter field, we obtain the interaction Hamiltonian

$$\begin{aligned} H = & (\kappa/2) [\bar{\gamma}_{\mu\nu} - \frac{1}{2} \partial_{\mu\nu} \gamma] T_{\mu\nu} + (\kappa^2/4) [T_{nn} \mathcal{A}^{-1} T_{nn} - 2T_{n\lambda}^s \mathcal{A}^{-1} T_{n\lambda}^s + T_{n\mu}^s \partial_\mu^s \partial_\nu^s \mathcal{A}^{-2} T_{n\nu}^s] \\ & + \kappa [\partial_\mu A_\nu - \frac{1}{2} \partial_{\mu\nu} \partial_\lambda A_\lambda] T_{\mu\nu} + (\kappa^2/2) \frac{1}{1 - \kappa [2\partial A_n + \partial_\lambda A_\lambda]} [(\partial A_p + \partial_p A_n - n_p \partial_\lambda A_\lambda) \partial_p V]^2 \end{aligned} \quad (4.15)$$

where

$$T_{\mu\nu} = \frac{1}{2} [\partial_\mu V \cdot \partial_\nu V - \delta_{\mu\nu} \{ (\partial_\lambda V)^2 + m^2 V^2 \}]. \quad (4.16)$$

The interaction Hamiltonian is then decomposed into the terms containing $\bar{\gamma}_{\mu\nu}$ and γ (whose commutation relation is given by (3.19) and the second equation of (3.20), $[\gamma, \gamma'] \doteq -4iD(x-x')$), the gravitational potential given by the second term of (4.15) and the terms containing A_μ . Using $-(\kappa/4)\gamma T_{\lambda\lambda}$ and the second term of (4.15) we obtained in [I] the equation of two-body problem in general relativity found by Einstein, Infeld and Hoffmann. As A_μ are arbitrary quantities, they will be unquantizable part of the gravitational field. Then, either putting them to zero or making use of the suitable canonical transformation, we can drop the terms containing A_μ . Thus, we are able to show that our formalism without coordinate condition gives the same descriptions as [I]. It seems to be interesting to compare the present formalism in this section with the canonical formalism given in § 3. Integrating by parts the first term of \mathfrak{H} given by (3.14) and using χ_μ^c equation (3.39), we can drop the interaction terms containing $\gamma_{n\mu}^s$.

Though we cannot drop the term containing γ_{nn} , it can be absorbed into the term containing γ . Therefore, also in § 3, γ_{nn} are not contained in the interaction terms.

§ 5. Concluding remarks

We have shown that we are able to formulate the quantum theory of the gravitational field without using the surplus coordinate condition which causes the mathematical difficulties, so as to give the same result as [I]. So we may conclude that the present formulation is a more reasonable procedure than the one with the coordinate condition.

In § 3, it is shown that the Dirac's canonical formalism is very useful for the system with the interaction Lagrangian with higher derivatives for which the unitary transformation connecting the free operator and the Heisenberg's one does not exist. Therefore, on the analogy of § 3, the quantization for the system with non-local action may be carried out consistently by using the Dirac P.b. instead of the usual one. Detailed accounts of this point will be given on another occasion.

The author wishes to express his sincere thanks to Prof. Y. Mimura for his kind interest in this work.

References

- 1) T. Kimura, *Prog. Theor. Phys.* **16** (1956), 157.
- 2) S. N. Gupta, *Proc. Phys. Soc.* **65** (1952), 161, 608; *Phys. Rev.* **96** (1954), 1683.
- 3) J. G. Valatin, *Danske Vidensk. Selsk. Mat.-fys. Medd.* **26** (1951) No. 13.
- 4) P. A. M. Dirac, *Can. Jour. Phys.* **33** (1955), 650.
- 5) S. Ozaki, *Prog. Theor. Phys.* **14** (1955), 511.
- 6) J. Schwinger, *Phys. Rev.* **74** (1948), 1439, (3-19).
- 7) Cf. G. Wentzel, *Quantum Theory of Fields* (1949), p. 207.
- 8) P. A. M. Dirac, *Can. Jour. Math.* **2** (1950), 129.
- 9) H. Umezawa and Y. Takahasi, *Prog. Theor. Phys.* **9** (1953), 14, 501.
Y. Katayama, *Prog. Theor. Phys.* **10** (1953), 129.
- 10) R. Utiyama, *Phys. Rev.* **101** (1956), 1597; *Prog. Theor. Phys.* **2** (1947), 38; **3** (1948), 14.

A Lattice Model of Liquid Helium, I

Takeo MATSUBARA and Hirotugu MATSUDA

Research Institute for Fundamental Physics, Kyoto University, Kyoto

Quantum Chemistry Laboratory, Chemistry Department, Kyoto University, Kyoto

(Received August 6, 1956)

A lattice model is proposed in order to explain the peculiar properties of liquid helium II. In the present model, liquid is regarded as a lattice composed of atoms and holes. Each atom can migrate by exchanging its position with an adjacent hole if the latter exists. In terms of operators which create or annihilate an atom at each lattice point, the kinetic and potential energies are defined so as to have a correct form in the limit of vanishing lattice spacing. Then it is proved that the grand partition function of our lattice liquid is equal to the partition function of a system of vector spins subjected to an external magnetic field, provided a proper identification is made between corresponding quantities of both systems. Various approximation methods used in the problem of ferromagnets are applicable and enable us to attain the results which explain most of the characteristic features of liquid helium.

§ 1. Introduction

In recent years, many attempts have been made to explain the peculiar properties of liquid helium II from the view-point of atomic theory.¹⁾ For the equilibrium properties, to which we shall confine ourselves in this paper, two quite different approaches have so far been adopted. One is the attack made from very low temperature side, looking for the ground state and low-lying excitations of liquid helium as a whole, and the other is the quantum statistical treatment of many Bose particles, aiming at elucidating the mystery of the λ -transition. As to the first approach, the concepts of phonon and roton excitations, originally introduced by Landau in somewhat phenomenological way,²⁾ are now not only successful in explaining the equilibrium as well as dynamical properties of liquid helium near absolute zero, but are also well scrutinized from the first principle of quantum mechanics.³⁾ On the other hand, the quantum statistical interpretation of the λ -transition seems far from being complete. Since F. London, the Bose-Einstein condensation of ideal Bose gas has been a leading idea to explain the λ -transition, and how to include the interatomic interaction into ideal Bose gas has been a main problem. Although many ingenious mathematical tricks have been invented to overcome the difficulties encountered in dealing with quantum statistical many body problems, most of the theories so far presented on this problem are essentially such ones that treat liquid helium as an imperfect Bose gas and accordingly fail to take account of the large atomic correlation in liquid phase.⁴⁾ This is the reason why the gas-like treatment of the λ -transition

cannot afford but results which reveal serious discrepancies when compared with the real phase transition of liquid helium. One may cite, for instance, the following points, of which all Bose-Einstein condensation theories could not give a satisfactory explanation:

- (a) The λ -transition is of the second kind.
- (b) The transition temperature T_λ decreases with increasing pressure p or density ρ , i.e.,

$$dT_\lambda/dp < 0 \text{ or } dT_\lambda/d\rho < 0.$$

- (c) The specific volume under constant pressure increases with decreasing temperature below λ point.

In order to settle these questions, it would be necessary to make a liquid-like treatment of many Boson system.

Beside these unsolved questions, furthermore, there remains another problem of how to reconcile the two approaches or the two concepts of phonon-roton excitations and Bose-Einstein condensation. A few attempts are reported on this point,⁵⁾ but the interrelation between the two different approaches does not seem to have received a good elucidation.

In these circumstances, it would not be meaningless to confine ourselves for a moment to a simple model which keeps the essential features of actual liquid helium and yet is easily dealt with, leading to a definite conclusion about the behaviors over the whole range of temperature. The aim of this paper is to work out such a model as is capable of clarifying all the peculiarities of liquid helium as simply as possible. In § 2 a lattice model is proposed for liquid helium and the Hamiltonian of the lattice liquid is set up. It turns out in § 3 that the grand partition function of the lattice liquid defined in § 2 is proved to be mathematically equivalent to the partition function of a system of vector spins and thus the problem is reduced to the study of the thermodynamical properties of a ferromagnet in an external magnetic field. In § 4 a molecular field approximation is adopted in order to examine the nature of the phase transition expected in both systems. In a crudest approximation (Weiss approximation), our model provides a second kind phase transition, transition temperature of which having a correct dependency on the density. In § 5 the nature of the states near absolute zero is studied by making use of spin-wave approximation in the problem of ferromagnets. This approximation affords us in a natural way the phonon excitation of lattice liquid.

§ 2. Lattice model of liquid helium

The lattice model for classical liquid is a well-known method which provides an adequate approximation for the purpose of taking into consideration the large atomic correlation in liquid phase. Each atom is assumed to occupy only discrete lattice points and the configurations of liquid are described by the distributions of atoms and holes among the lattice points. By choosing a proper magnitude of lattice constant, and excluding the multiple occupations of atoms on each lattice point, one can account for the effect of the strong repulsion between atoms.

In extending this idea of lattice model to quantum liquid such as liquid helium, it is of primary importance to take account of the effect of the zero point motion of the atoms.⁽⁴⁾ Since considerable increment in kinetic energy will be caused by localizing each atom in a lattice configuration, it must be reduced by mixing various configurations. This means that, in the lattice representation on which our lattice model rests, the kinetic energy has large non-diagonal elements such as to produce transitions among various lattice configurations. Taking consideration of this effect of the kinetic energy on one hand, and of the excluding effect of the strong repulsion between atoms on the other hand, we shall construct a lattice model for liquid helium in the following way.

For simplicity of presentation we take a simple cubic lattice with a lattice constant d . It is convenient to adopt the scheme of the second quantization, and define operators a_i^* and a_i which creates or annihilates an atom at the i -th lattice point. For Bose particles, we assume the commutation relations

$$[a_i^*, a_j^*]_- = [a_i, a_j]_- = [a_i, a_j^*]_- = 0 \quad (1)$$

for $i \neq j$, and to exclude the multiple occupation of atoms at each point, we impose for $i = j$ the restrictions

$$[a_i^*, a_i^*]_+ = [a_i, a_i]_+ = 0, \quad [a_i, a_i^*]_+ = 1. \quad (2)$$

In other words, the operators with different lattice indices are commutable to each other and with respect to the same lattice point they have a character like Fermi particles. One might ask whether such operators as satisfying (1) and (2) simultaneously could exist, but a momentary reflection would make one recall that they are nothing but a set of algebra equivalent to the spin with magnitude of $1/2$.

Now first we shall set up the potential energy of the lattice liquid. We consider here the following potential energy v between two atoms;

$$v = \infty \quad \text{if the two atoms occupy the same lattice point (this case is forbidden by (2)),}$$

$$v = -v_0 \quad \text{if the two atoms are nearest neighbours}$$

$$\text{and } v = 0 \quad \text{otherwise.}$$

Then the total potential energy of the system may be written as

$$\Phi = -v_0 \sum_{\langle ij \rangle} a_i^* a_j^* a_j a_i, \quad (3)$$

where $\langle ij \rangle$ means to take the summation over all the nearest neighbouring pair points. $a_i^* a_i$ represents the number of atoms occupying the i -th point, and therefore the total number of atoms is given by

$$\sum_i a_i^* a_i = N_0. \quad (4)$$

As to the kinetic energy, we shall make an assumption that each atom can make a transition only to one of nearest neighbour sites when it is vacant. Such a transition is evidently generated by the operator of the form $a_i^* a_j$, where i and j are nearest neighbour to each other, and the non-diagonal part of the kinetic energy will be a sum of the operators of

this form. In effect, we propose as the total kinetic energy of the system

$$K = (\hbar^2/2md^2) \sum_{\langle ij \rangle} (a_i^* - a_j^*) (a_i - a_j), \quad (5)$$

which is suggested by inspecting a limit of the kinetic energy of continuum

$$\frac{\hbar^2}{2m} \int \frac{\partial \psi^{*}}{\partial \mathbf{x}} \frac{\partial \psi}{\partial \mathbf{x}} d\mathbf{x},$$

where m is the mass of the particles and $\psi^{*}(\mathbf{x})$, $\psi(\mathbf{x})$ are quantized wave functions. Making Fourier transforms of a_i^* and a_i by the relations

$$\begin{aligned} a_k^* &= N^{-1/2} \sum_j \exp(-ikr_j) a_j^*, \\ a_k &= N^{-1/2} \sum_j \exp(ikr_j) a_j, \end{aligned} \quad (6)$$

where N is the total number of lattice points and \mathbf{r}_i the positional vector of the i -th point, and on expressing K in terms of (6) one gets

$$K \cong \sum_k (\hbar^2 k^2 / 2m) a_k^* a_k, \quad (7)$$

provided $kd \ll 1$. This agrees precisely with the kinetic energy of free particles and thus (5) is proved to have a correct form for long waves.

After all, we start with a Hamiltonian of the lattice liquid given by

$$\mathcal{H}_L = (\hbar^2/2md^2) \sum_{\langle ij \rangle} (a_i^* - a_j^*) (a_i - a_j) - v_0 \sum_{\langle ij \rangle} a_i^* a_i a_j^* a_j. \quad (8)$$

To know the thermodynamical properties of our lattice liquid, it is most convenient to construct the grand partition function defined by

$$\Xi_L = \text{Trace}[\exp(-\alpha \sum_i a_i^* a_i - \beta \mathcal{H}_L)], \quad (9)$$

from which all thermodynamical quantities are readily calculated. For instance, using characteristic function $\Psi = \log \Xi_L$,

$$\text{the number of atoms} \quad N_0 = -(\partial \Psi / \partial \alpha)_\beta, \quad (10a)$$

$$\text{internal energy} \quad E = -(\partial \Psi / \partial \beta)_\alpha, \quad (10b)$$

$$\text{pressure} \quad pV/kT = \Psi \text{ etc.} \quad (10c)$$

On the other hand, as was already pointed out, the operators satisfying (1) and (2) are equivalent to a set of N spins each of which is localized at each lattice point and has the magnitude of $1/2$. Indeed, for $S=1/2$ the spin operators

$$S_{j\pm} = S_{jx} \pm iS_{jy} \quad (11)$$

satisfy the following commutation relations:

$$[S_{i+}, S_{j-}]_- = [S_{i-}, S_{j-}]_- = [S_{i-}, S_{j+}]_- = 0 \quad (i \neq j) \quad (1')$$

$$[S_{i+}, S_{i+}]_+ = [S_{i-}, S_{i-}]_+ = 0, \quad [S_{i-}, S_{i+}]_+ = 1. \quad (2')$$

These are the same as (1) and (2). Furthermore, as the z -component of spin S_{iz} can be expressed in terms of S_{i+} and S_{i-} as

$$S_{iz} = S_{i+} S_{i-} - 1/2, \quad (12)$$

there should exist a certain system of spins equivalent to the lattice liquid with a Hamiltonian (8) by the correspondence $a_i^* \longleftrightarrow S_{i+}$, $a_i \longleftrightarrow S_{i-}$. Let us study such a system of spins in the next section.

§ 3. A spin system equivalent to the lattice model

Let us consider a spin system whose Hamiltonian is given by

$$\mathcal{H}_F = -J \sum_{\langle ij \rangle} (S_{ix} S_{jx} + S_{iy} S_{jy}) - J' \sum_{\langle ij \rangle} S_{iz} S_{jz} - H \sum_i S_{iz}. \quad (13)$$

\mathcal{H}_F represents the energy of spins which interact with an anisotropic exchange coupling and are subjected to an external magnetic field. Making use of (11) and (12), (13) can be easily rewritten as

$$\begin{aligned} \mathcal{H}_F = & J/2 \cdot \sum_{\langle ij \rangle} (S_{i+} - S_{j+}) (S_{i-} - S_{j-}) - J' \sum_{\langle ij \rangle} S_{i+} S_{i-} S_{j+} S_{j-} \\ & - (H + z/2 \cdot J - z/2 \cdot J') \sum_i S_{i+} S_{i-} + N/2 \cdot (H - 1/4 \cdot z J'), \end{aligned} \quad (14)$$

where z is the number of nearest neighbours. Now let the partition function of this spin system be

$$\Xi_F = \text{Trace} [\exp(-\beta \mathcal{H}_F)]. \quad (15)$$

Then, comparing (8) and (9) with (14) and (15), it is readily realized that Ξ_L and Ξ_F are equal to each other except for a multiplication constant, namely

$$\Xi_L = C \Xi_F, \quad (16)$$

if the following identification is made:

$$\hbar^2/2md^2 = J/2, \quad v_0 = J', \quad (17)$$

$$\alpha = -\beta \{H + (z/2) J - (z/2) J'\}, \quad C = \exp[(N/2)\beta \{H - (z/4) J'\}].$$

Therefore, if we know the thermodynamical properties of our spin system, we can presume the behaviours of our lattice liquid by simple translation, and vice versa.

The relations between the physical quantities of both systems are straightforwardly found. First, the density of the lattice liquid is connected with the magnitude of the magnetization of the spin system along z -axis. For, by virtue of (10a), (16), (17),

$$\begin{aligned} N_0 &= -\frac{\partial}{\partial \alpha} (\log \Xi_L) \\ &= \frac{\partial}{\partial (\beta H)} [\log C + \log \Xi_F] \\ &= N/2 + \langle \sum_i S_{iz} \rangle. \end{aligned}$$

Or, if we denote the density of the lattice liquid by $N_0/N=\mu$, and the mean value of the z -component of a spin by $\langle S_z \rangle = \zeta$, then it holds that

$$2\rho=1+2\zeta. \tag{18}$$

It is natural to assume $\rho > 1/2$ for liquid phase and so we may confine ourselves to the case $\zeta > 0$. Next, by (17) the chemical potential of the lattice liquid $\mu = -\alpha/\beta$ corresponds to the external magnetic field apart from an additive constant. And the pressure p of the lattice liquid is identified with a linear combination of free energy per one spin f and external magnetic field measured in proper energy unit :

$$pd^3 = -f + 1/2 \cdot (H - z/4 \cdot J'), \tag{19}$$

where $f = -(1/\beta N) \log \Xi_F$ and we put the volume $V = Nd^3$. In Table I we summarize the relations between the corresponding quantities of both systems.⁷⁾

It is noted that the kinetic energy of the lattice liquid corresponds to the exchange energy of spins which prefer to direct perpendicularly to the z -axis, if external magnetic field is absent, on account of anisotropic exchange coupling. In other words, to maximize

Table I
Identification of corresponding quantities

| Lattice liquid | | system of ferromagnets |
|-----------------------|---|------------------------------------|
| volume/ d^3 | = | number of spins |
| number density ρ | = | $\zeta + 1/2$ |
| chemical pot. μ | = | $H + z/2 \cdot (J - J')$ |
| pressure pd^3 | = | $-f + 1/2 \cdot H - 1/8 \cdot zJ'$ |

the negative of exchange energy by making as many spins as possible point to one direction in the xy plane corresponds to minimize the kinetic energy of the lattice liquid by making the localization of each atom as small as possible. And when the spin system becomes to magnetize spontaneously at sufficiently low temperatures, there appears a long range order by which the direction of any two spins, however distant, are correlated to each other. Corresponding to this phenomena of long range order in ferromagnets, there would appear in the lattice liquid a kind of "long range order of momentum", which might have a bearing on superfluidity.⁸⁾

§ 4. Phase transition

In the preceding section, we have found that there exists an intimate connection between the thermodynamical properties of a system of ferromagnets and that of lattice liquid. Therefore, one can expect that the appearance of ferromagnetic ordering in one system will be related to the occurrence of the λ -transition or of superfluidity in the other. In order to examine this relation, it is more convenient to start with the ferromagnetic

system, because the physical phenomena involved in this system are more familiar and more easily understood by intuition. First we shall classify two cases according as $J' > J$ or $J > J'$.

The case of $J' > J$.

This is the case where the exchange coupling in the z direction is stronger than that in the x or y direction, and hence at sufficiently low temperatures a spontaneous magnetization appears along the z -axis. This implies that in the lattice liquid there appears an order of the distribution of atoms and holes such that atoms and holes congregate separately to each other. This case should rather be regarded as a case where solid-liquid or liquid-gas condensation occurs. To exclude such a phase separation, we must impose a condition

$$J \geq J'$$

or

$$\hbar^2/md^2 \geq v_0. \quad (20)$$

(20) should be considered as a condition that the lattice system is a homogeneous "quantum liquid" down to absolute zero.

The case of $J > J'$.

In this case a spin ordering in the xy plane will exhibit at sufficiently low temperatures. To see what happens in the lattice liquid, we shall first solve the problem of ferromagnetism by means of molecular field approximation, and then translate the results into the languages of the lattice liquid.

Replacing the exchange interaction between neighbouring spins by molecular field, the one spin Hamiltonian is written as

$$\mathcal{H}_1 = -zJ\hat{\xi}S_x - (H + zJ'\zeta)S_z = -aS_x - bS_z, \quad (21)$$

where

$$a = zJ\hat{\xi}, \quad b = H + zJ'\zeta, \quad (22)$$

and $\hat{\xi}$ and ζ are statistical averages of S_x and S_z respectively, defined by

$$\begin{aligned} \hat{\xi} &= \frac{\partial}{\partial(\beta a)} \log \text{Trace} [\exp(-\beta\mathcal{H}_1)], \\ \zeta &= \frac{\partial}{\partial(\beta b)} \log \text{Trace} [\exp(-\beta\mathcal{H}_1)]. \end{aligned} \quad (23)$$

The eigenvalues of (21) are easily obtained as $\pm (1/2)\sqrt{a^2 + b^2}$, and by noting that

$$\text{Trace} [\exp(-\beta\mathcal{H}_1)] = 2 \cosh \frac{\beta}{2} \sqrt{a^2 + b^2}$$

(23) are reduced to

$$\hat{\xi} = \frac{a}{2\sqrt{a^2 + b^2}} \tanh \frac{\beta}{2} \sqrt{a^2 + b^2}, \quad (24)$$

$$\zeta = \frac{b}{2\sqrt{a^2+b^2}} \tanh \frac{\beta}{2} \sqrt{a^2+b^2}. \quad (25)$$

Now there appear two cases :

$$(i) \quad a=0, \text{ i.e. } \hat{\xi}=0.$$

This is a trivial solution of (24). In this case the spontaneous magnetization is absent and from (25) the magnetization along the z -axis is given by

$$\zeta = \frac{1}{2} \tanh \frac{\beta}{2} b$$

or

$$2\zeta = \tanh \left(\frac{H+zJ'\zeta}{2kT} \right). \quad (26)$$

As is proved shortly, this state is stable over all the range of temperature when H exceeds a critical field H_c and is stable only above a critical temperature T_c when $H < H_c$.

$$(ii) \quad a \neq 0.$$

From (24) it follows that

$$\frac{2}{zJ} \sqrt{a^2+b^2} = \tanh \frac{\beta}{2} \sqrt{a^2+b^2},$$

or putting

$$X = (2/zJ) \sqrt{a^2+b^2}, \quad T_0 = zJ/4k, \quad (27)$$

$$X = \tanh (T_0/T) X. \quad (28)$$

The equation (28) has a non-zero solution $0 \leq X(T) \leq 1$ when $T < T_0$, in addition to a trivial solution $X=0$. On the other hand, combining (25) with (24) one gets

$$\zeta = b/zJ = (H+zJ'\zeta)/zJ$$

and thence

$$\zeta = H/z(J-J'), \quad (29)$$

$$\hat{\xi}^2 + \zeta^2 = (1/4) X^2(T). \quad (30)$$

When H is given, ζ is a constant (independent of temperature) and has to satisfy an inequality

$$\zeta \leq (1/2) X(T) \quad \text{or} \quad 2H/z(J-J') \leq X(T) \leq 1,$$

from which the critical field H_c is determined :

$$H_c = (z/2) \cdot (J-J'). \quad (31)$$

For $H > H_c$, $\hat{\xi}$ cannot be finite and we should go back to the case (i) where $\hat{\xi}=0$. For $H < H_c$, ζ is a constant given by (29), whereas $\hat{\xi}$ decreases with increasing temperatures according to the equations (28) and (30) and lastly vanishes at the critical temperature

T_c determined by

$$2\zeta = 2H/\chi(J-J') = X(T_c). \quad (32)$$

Above the critical temperature $T > T_c$, ζ remains vanishing and again we have the case (i). We summarize these results in Figs. 1 and 2, in which ζ is plotted as a function of T (under constant field H) or of H (under constant T). The physical meaning of these results will be quite obvious. We are dealing with an assembly of spins, each of which is subjected to two competitive forces. One is the external magnetic field which

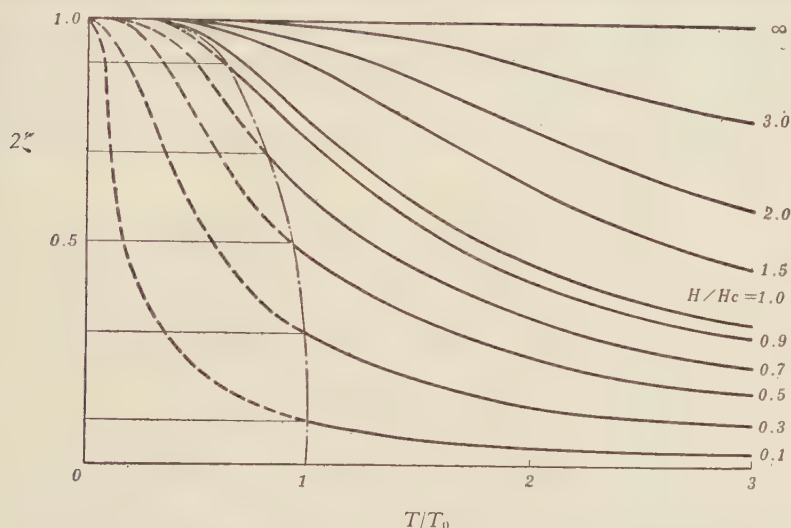


Fig. 1. Magnetization versus temperature curves for constant magnetic fields.

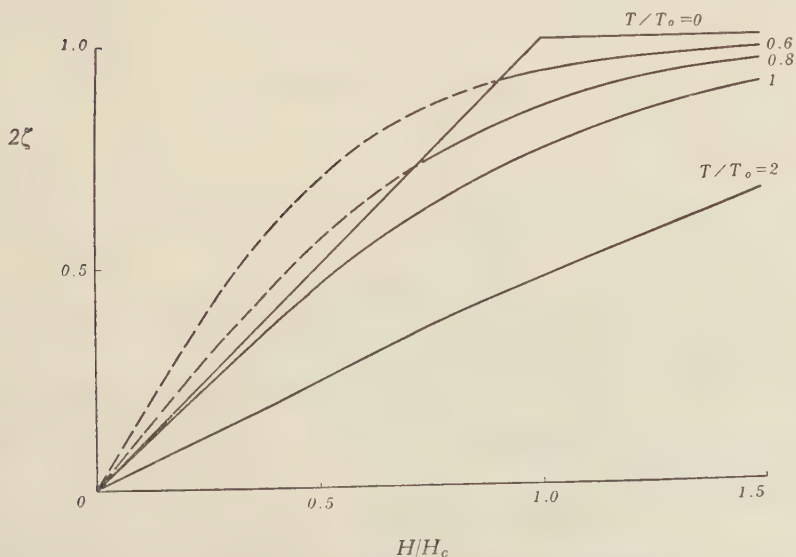


Fig. 2. Magnetization versus field curves for constant temperatures

turns the spins to the z -direction and the other is the exchange force which acts so as to put the spins into order in the plane perpendicular to the z -axis. When the external field overwhelms the exchange force, the ferromagnetic ordering is impossible, whereas a cooperative ordering of spins is established when the exchange force surpasses the decoupling effect of the external field and thermal agitation. Once the ferromagnetic ordering in the xy plane is set up, it becomes difficult to magnetize the spin system along the direction of the field.

Now let us see what kind of phase change occurs in the lattice liquid. As the density of the lattice liquid is related to the magnetization along the z -axis and chemical potential μ to magnetic field H , the variation of μ with temperature will be known by seeking for the cross points of the constant field curves in Fig. 1 with a horizontal line corresponding to a given density. The results are shown in Fig. 3, from which one can draw out the following conclusions:

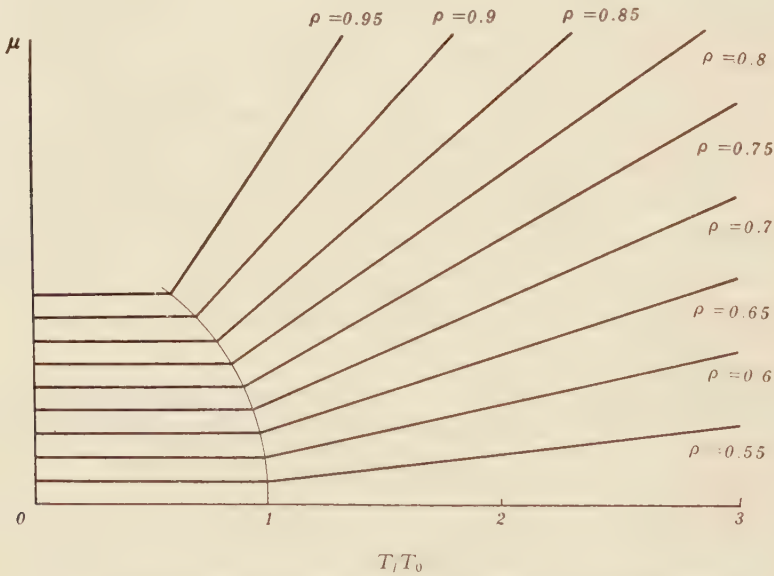


Fig. 3. Chemical potential μ versus temperature curves for various densities.

- (1) A phase transition occurs at a critical temperature T_c .
This is of the second kind, because the chemical potential reveals a kink at this point.
- (2) The critical point T_c decreases with increasing density.

Making use of (28) and (32), the density dependence of T_c is given by

$$(2\rho - 1) = \tanh (T_0/T_c) (2\rho - 1), \tag{33}$$

where

$$T_0 = zJ/4k = z\hbar^2/4md^2k.$$

If we identify this transition with the λ -transition of liquid helium, the equation

(33) yields a fairly excellent agreement between calculated and observed transition temperatures. For instance, if we take m as the true mass of He^4 (i.e. $m=6.68 \times 10^{-24}$ gr.) and choose a lattice constant $d=3 \times 10^{-8}$ cm with $z=6$, then

$$T_0=2^0 \text{ K.}$$

For several values of mass density $\rho_M=m/d^3 \cdot \rho$, the calculated T_λ values are tabulated in Table II together with the observed values.

Table II

| ρ_M | ρ | T_λ (calculated) | T_λ (observed) |
|-------------------------|--------|--------------------------|------------------------|
| 0.15 gr/cm ³ | 0.61 | 1.98 °K | 2.15 °K |
| 0.16 | 0.65 | 1.95 | 2.05 |
| 0.17 | 0.69 | 1.89 | 1.9 |
| 0.175 | 0.71 | 1.83 | 1.8 |

All other thermodynamical quantities will be easily evaluated with the aid of Table I in the preceding section, if the free energy of the spin system is known. The free energy per one spin is in the present approximation evaluated as

$$f=\frac{z}{2} J \zeta^2+\frac{z}{2} J' \zeta^2-k T \log \left[2 \cosh \frac{2 T_0}{T} \sqrt{\zeta^2+\frac{(H+z J' \zeta)^2}{(z J)^2}} \right]. \quad (34)$$

The condition $\partial f / \partial \zeta=0$ and $\partial f / \partial \zeta'=0$ give the equation (24) and (25) respectively. One cannot, however, expect a quantitative agreement between theory and experiment because of the neglect of short range order. In the next paper we shall adopt a more refined approximation to calculate various thermodynamical quantities and to compare them with observations.

§ 5. Phonon and spin wave

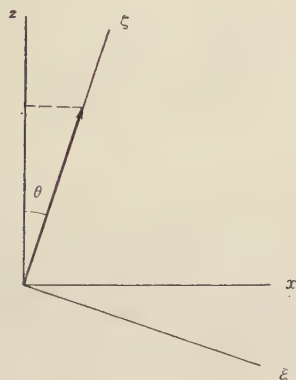


Fig. 4.

It is well known that the state of ferromagnets near absolute zero can be correctly described by the spin-wave theory and that of liquid helium can be dealt with as an assembly of phonons. Therefore, these two kinds of excitations are expected to relate to each other in our model.

In order to derive the spin wave Hamiltonian from (13), we assume that the magnetization at absolute zero makes an angle θ with z -axis (the direction of external magnetic field, see Fig. 4), and perform a transformation of the spin coordinates

$$S_x=S_\xi \cos \theta+S_\zeta \sin \theta$$

$$S_y = S_\eta \quad (35)$$

$$S_z = -S_\xi \sin \theta + S_\zeta \cos \theta,$$

where ζ -axis is the direction of the magnetization. Then \mathcal{H}_F becomes

$$\begin{aligned} \mathcal{H}_F = & - \sum_{\langle ij \rangle} [(J \cos^2 \theta + J' \sin^2 \theta) S_{i\zeta} S_{j\zeta} + J S_{i\eta} S_{j\eta} + (J \sin^2 \theta + J' \cos^2 \theta) S_{i\xi} S_{j\xi} \\ & + (J - J') \sin \theta \cos \theta (S_{i\zeta} S_{j\zeta} + S_{i\xi} S_{j\xi}) + H \sin \theta \sum_i S_{i\zeta} - H \cos \theta \sum_i S_{i\xi}. \end{aligned} \quad (36)$$

Now we shall assume that in the ground state $S_{i\zeta}$'s take nearly constant value $S = 1/2$ and introduce, after Anderson,⁽⁹⁾ the normal coordinates of spin waves by the relations

$$S_{j\xi} = (2N)^{-1/2} \sum_k \exp(i\mathbf{k}\mathbf{r}_j) Q_k, \quad S_{j\eta} = (2N)^{-1/2} \sum_k \exp(-i\mathbf{k}\mathbf{r}_j) P_k, \quad (37)$$

$$S_{j\zeta} \simeq 1/2 - (S_{j\xi}^2 + S_{j\eta}^2 - 1/2). \quad (38)$$

The commutation relations for P_k and Q_k are

$$[Q_k, Q_{k'}]_- = [P_k, P_{k'}]_- = 0, \quad (39)$$

$$[Q_k, P_{k'}]_- = (2i/N) \sum_j \exp[i(\mathbf{k} - \mathbf{k}')\mathbf{r}_j] S_{j\zeta} \quad (40)$$

the last of which may be approximated as

$$[Q_k, P_{k'}]_- = i\delta_{kk'} \quad (41)$$

so long as $S_{j\zeta}$ is considered as a constant. θ should be determined in such a way as to make the linear term in $S_{j\xi}$ and $S_{j\eta}$ vanish in the transformed Hamiltonian⁽¹⁰⁾. Putting (37) and (38) into (36), and neglecting terms higher than square in P_k and Q_k , we obtain the equation to determine θ :

$$H \sin \theta - z/2 \cdot (J - J') \sin \theta \cos \theta = 0 \quad (42)$$

and the spin wave Hamiltonian becomes

$$\begin{aligned} \mathcal{H}_{spin \ wave} \simeq & -3/4 \cdot zNJ - (1/4) zN(J - J') \cos^2 \theta \\ & + z/2 \cdot \sum_k [J - \{J + (J' - J) \sin^2 \theta\} \gamma_k] Q_k Q_{-k} + z/2 J \cdot \sum_k (1 - \gamma_k) P_k P_{-k}, \end{aligned} \quad (43)$$

where

$$\gamma_k = 1/z \cdot \sum_j \exp(i\mathbf{k}\mathbf{r}_j). \quad (44)$$

The solution of (42) are obviously

$$\sin \theta = 0 \quad \text{or} \quad \cos \theta = \frac{2H}{z(J - J')},$$

and the latter is the desired one for weak field. The procedures of diagonalizing $\mathcal{H}_{spin \ wave}$ are well known and need not be described here. The energy of the assembly of spin waves is finally given by

$$\mathcal{H}_{spin \ wave} = C + \sum_k \hbar \omega_k \left(n_k + \frac{1}{2} \right), \quad (45)$$

where C is an additive constant

$$C = - (3/4) z N J - (1/4) z N (J - J') \cos^2 \theta,$$

and the energy of spin wave with wave vector \mathbf{k} , $\hbar\omega_k$, is expressed as

$$\hbar\omega_k = z/2 \cdot \sqrt{J(1-\gamma_k) \{J - (J \cos^2 \theta + J' \sin^2 \theta) \gamma_k\}}. \quad (46)$$

At very low temperatures, important contributions to free energy will come from long spin waves for which $kd \ll 1$ and γ_k may be approximated by

$$\gamma_k = 1 - 1/z \cdot (kd)^2 + \dots,$$

and hence the energy spectrum becomes

$$\hbar\omega_k \simeq (z/2) \sin \theta \sqrt{J(J-J')/z \cdot kd}, \quad (47)$$

which is linear in k just as the phonon spectrum is. This result suggests itself that the free energy of our ferromagnets as well as our lattice liquid will be proportional to T^4 and the specific heat to T^3 at very low temperatures, as actually observed in liquid helium. It would be interesting to calculate the speed of sound wave from the phonon spectrum and to compare the result with observation. The speed of sound wave c is given by

$$\begin{aligned} c &= \frac{\hbar\omega_k}{\hbar k} = \frac{z}{2\hbar} \sin \theta \sqrt{\frac{J(J-J')}{z}} d \\ &= \frac{\hbar}{md} \frac{\sqrt{z}}{2} \sin \theta \left(1 - \frac{md^2 v_0^2}{\hbar^2}\right)^{1/2}. \end{aligned} \quad (48)$$

Putting $J' = v_0 = 0$ and $\sin \theta \simeq 1$ for a moment and using the following numerical values of constants

$$m = 6.68 \times 10^{-24} \text{ gr.}$$

$$d = 3 \times 10^{-8} \text{ cm,} \quad z = 6,$$

one gets

$$c \sim 200 \text{ m/sec}$$

which is in excellent accord with the observed value 250 m/sec.

§ 6. Conclusions

We have proposed a lattice model for liquid helium and studied the equilibrium properties of our lattice liquid. In the crudest approximation we have obtained the results which are capable of explaining, at least qualitatively, the main characteristic of the transition of liquid helium. In order to proceed to a more quantitative discussion, however, it is necessary to employ a more refined approximation in evaluating the grand partition function of the system. In the next paper we shall undertake this refinement of calculations and present a detailed comparison between the observations and what our model provides.

References

- 1) As a review, for instance, see R. B. Dingle, *Advance in Physics* **1** (1952), 1.
- 2) L. Landau, *J. Phys. USSR* **5**(1941), 71.
- 3) R. P. Feynman, *Phys. Rev.* **91**(1953), 1301 ; **94** (1954), 262.
C. G. Kuper, *Proc. Roy. Soc. A* **223** (1955), 223.
- 4) F. London, *Phys. Rev.* **54** (1938), 947.
L. L. Schiff, *Phys. Rev.* **59** (1941), 751, 758 ; **60** (1941), 362.
T. Matsubara, *Prog. Theor. Phys.* **6** (1951), 714.
R. P. Feynman, *Phys. Rev.* **91** (1953), 1291.
G. V. Chester, *Phys. Rev.* **94** (1954), 246.
De ter Haar, *Phys. Rev.* **95** (1954), 895.
R. Kikuchi, *Phys. Rev.* **96** (1954), 563.
S. T. Butler & M. H. Friedman, *Phys. Rev.* **98** (1955), 287.
G. V. Chester, *Phys. Rev.* **100** (1955), 455.
- 5) W. Brenig, *Zeit. f. Phys.* **144** (1956), 488.
- 6) S. Koide, *J. Phys. Soc. Jap.* **11** (1956), 7.
Also see, I. Prigogine et J. Philippot, *Physica* **18** (1952), 227, 235.
J. de Boer, *Physica* **20** (1954), 655.
E. G. D. Cohen, J. de Boer & Z. W. Salsbury, *Physica* **21** (1955), 1937.
- 7) A similar correspondence was pointed out by Lee and Yang.
T. D. Lee & C. N. Yang, *Phys. Rev.* **87** (1952), 410.
- 8) I. M. Blatt, S. T. Butler & M. R. Schafroth, *Phys. Rev.* **100** (1955), 481.
- 9) P. W. Anderson, *Phys. Rev.* **86** (1952), 694.
- 10) J. Kanamori & K. Yosida, *Prog. Theor. Phys.* **14** (1955), 423.

Some Consequences of the Compound Hypothesis for Elementary Particles

Ken-iti MATUMOTO

Institute of Theoretical Physics, Nagoya University, Nagoya

(Received August 31, 1956)

It is the purpose of the present paper to derive some consequences of the hypothesis proposed recently by Sakata for the models of pions and of heavy unstable particles. On the basis of Sakata's model, we propose a semi-empirical mass formula for these particles and obtain some interesting results from this formula.

§ 1. Introduction

The knowledge of the properties of hyperons and heavy mesons has grown quickly in the last few years. These properties (especially associated production, abundant production, long lifetime and cascade decay) have been fairly well explained by the theory of Nishijima and Gell-Mann¹⁾. In their theory they introduced a new quantum number, "Strangeness", and made the well-known assignment of the isotopic spin (I) and of the strangeness (S) for each particle. The present experimental results may be stated as follows: the strong interactions conserve I and S and the weak interactions change the magnitude of S by ± 1 .

D'Espagnat and Prentki²⁾ attempted to understand the success of the Nishijima and Gell-Mann theory on the basis of the present quantum field theory. They formulated the Nishijima-Gell-Mann-model as follows: the strong interaction Lagrangian is invariant under isobaric rotations as well as isobaric reflections. Furthermore they noticed that in the strong interactions not only the fermion number is conserved but also the number of isofermions.

On the other hand, it has been attempted by Goldhaber³⁾, Markov⁴⁾ and Sakata⁵⁾ to understand the success of the Nishijima and Gell-Mann theory from substantial points of view. In Goldhaber's model, nucleons, pions and θ mesons are the fundamental particles and the hyperons and other heavy mesons are the composite particles of them. Markov assumed that nucleons and hyperons are the fundamental particles and pions and heavy mesons are composed from them.

According to Sakata's idea, only the nucleon (N) and Λ particle (Λ) are supposed to be the fundamental particles and pions and other heavy unstable particles are regarded as the composite particles which are composed of suitable numbers of N , Λ and their anti-particles (\bar{N} , $\bar{\Lambda}$) as is indicated in Table 1.

Thus, if we assign to the N , Λ the isotopic spin $1/2$, 0 and the strangeness 0 , -1 respectively, we can get automatically the composite particles having the various

magnitude of I and S . Among them there are the particles which are unstable for the strong interactions as well as the particles which are stable for them.

Table 1. Sakata's model

| particles | composition |
|----------------|---------------------------|
| π | $(N\bar{N})$ |
| θ, τ | $(N\bar{\Lambda})$ |
| Σ | $(N\bar{N}\Lambda)$ |
| Ξ | $(\bar{N}\Lambda\Lambda)$ |

According to this model, the ordinary charge independent strong interactions between nucleon family and pion family are reduced to the assumed charge independent forces acting between the constituents of the composite particles⁽⁶⁾.

Therefore, with this model, we can reproduce all the essential points of the Nishijima-Gell-Mann-model for the observed particles. Furthermore, from the consideration of the existence of the particles which are unstable for the strong interactions we may have some further insight into the various high energy phenomena.

In addition, referring to the applicability of the present quantum field theory (especially that of the meson theory), we might expect that the compound effect plays the role of a high energy cut-off and that this model therefore raises the magnitude of the critical energy above which the so-called "Ghost" state appears.

On the basis of Sakata's model, in § 2, we propose the semi-empirical mass formula for the particles having various magnitudes of I and S . In § 3, we investigate the consequences which are obtained from the mass formula.

§ 2. Mass formula

On the basis of Sakata's model, we propose the following mass formula for the particles in their "ground state" having various magnitudes of I and S (except for the particle with $I=0$ and $S=0$):

$$\begin{aligned}
 M = & 1840 \times (N_N + N_{\bar{N}}) + 2180 \times (N_{\Lambda} + N_{\bar{\Lambda}}) \\
 & - 3400 \times (N_{N\bar{N}} - N_{NN} - N_{\bar{N}\bar{N}}) - 3050 \times (N_{N\bar{\Lambda}} + N_{\bar{N}\Lambda} - N_{N\Lambda} - N_{\bar{N}\bar{\Lambda}}) \\
 & - 2500 \times (-N_{\Lambda\Lambda} - N_{\bar{\Lambda}\bar{\Lambda}})
 \end{aligned} \quad (1)$$

where $N_N, N_{\bar{N}}, N_{\Lambda}, N_{\bar{\Lambda}}$, denote the numbers of $N, \bar{N}, \Lambda, \bar{\Lambda}$ in the composite particles and $N_{N\bar{N}}, N_{NN}, N_{\bar{N}\bar{N}}, N_{N\bar{\Lambda}}, N_{\bar{N}\Lambda}, N_{N\Lambda}, N_{\bar{N}\bar{\Lambda}}, N_{\Lambda\Lambda}, N_{\bar{\Lambda}\bar{\Lambda}}$ denote the numbers of the pairs of $N\bar{N}, NN, \bar{N}\bar{N}, N\bar{\Lambda}, \bar{N}\Lambda, N\Lambda, \bar{N}\bar{\Lambda}, \Lambda\Lambda, \bar{\Lambda}\bar{\Lambda}$ in the composite particles respectively. We use the electron mass unit in this paper.

In (1) we have assumed the following:

(i) the total mass defect of the many-body system equals the sum of the two-body mass defects of all the bonds involved,

$$\begin{aligned}
 \text{(ii)* } \left. \begin{aligned} [N\bar{N}] &= -[NN] = -[\bar{N}\bar{N}] & (=m_N + m_{\bar{N}} - m_{\pi} = 3400), \\ [N\bar{\Lambda}] (= [\bar{N}\Lambda]) &= -[\Lambda N] = -[\bar{\Lambda}\bar{N}] & (=m_N + m_{\bar{\Lambda}} - m_{\theta} (=m_{\bar{\theta}}) = 3050), \\ [\Lambda\Lambda] &= [\bar{\Lambda}\bar{\Lambda}] & (=m_N + 2m_{\Lambda} - m_{\Xi} - 2[\bar{N}\Lambda] = -2500), \end{aligned} \right\} \quad (2)
 \end{aligned}$$

* $(N\bar{N})_{I=0}$ and $(\Lambda\bar{\Lambda})$ virtually go over into each other through the intermediary of the strong

where $[AB]$ denotes the two-body mass defect due to the bond $A-B$.

These assumptions are based on the following points:

- (i) the stable particles of fermion number 2 do not exist;
- (ii) the invariance under the charge conjugation;
- (iii) assuming $[N\bar{N}] = -[NN]$ and $[\bar{N}\Lambda] = -[N\Lambda]$ and employing the values $[N\bar{N}] = m_N + m_{\bar{N}} - m_\pi = 3400$ and $[\bar{N}\Lambda] = m_{\bar{N}} + m_\Lambda - m_{\bar{p}} = 3050$ we find the following relations

$$m_{\Sigma=(N\bar{N}\Lambda)} = m_\Lambda + m_N + m_{\bar{N}} - [\Lambda N] - [\Lambda \bar{N}] - [N\bar{N}], \quad (3)$$

$$m_{F_0^{3/2}=(N\bar{N}N)} = 2m_N + m_{\bar{N}} - [NN] - 2[N\bar{N}], \quad (4)$$

within an error of ~ 200 which is very small compared to m_N , m_Λ , $[N\bar{N}]$ and $[\bar{N}\Lambda]$. Where $F_0^{3/2}$ is the fermion having $I=3/2$ and $S=0$ and we identify it as the resonance level of $I=3/2$ and spin $=3/2$ which comes into play in $\pi-N$ scattering in the region of the incident pion energy of 190 Mev. So the spin of $F_0^{3/2}$ is supposed to be $3/2$.

For the three-body particles, (1) reproduces the mass values within an error of ~ 200 . For the four- or five-body particles of course, the mass values obtained from (2) will be subject to larger errors.

§ 3. Consequences from the mass formula

For the particles having various magnitude of I and S , we show, in Table 2, the composition given by Sakata and the mass values deduced from the mass formula.

The characteristic results obtained are as follows.

i) On the "new" particles which may be found in future

"New" particles which may be found in future will not be so numerous. The mass formulas show that there are bosons B_2^0 , B_2^1 , B_{-2}^0 and B_{-2}^1 with mass of $1700 \sim 1800$ and fermions F_{-3}^0 and F_{-3}^1 with mass of about 2800. These particles are stable energetically for the strong interactions and so they may have the chance of observation.

Since we are considering four- or five-body particles, errors of the mass formula, (2), are probably not small. But we may say that the "new" particles, if discovered, may reasonably be identified as the above particles.

In this case, these particles will decay through the weak interactions as follows,

interactions. This is not the case between $(NN)_{I=0}$ and $(\Lambda\Lambda)$. This contribution to the energy is unknown. Non evidence of the existence of the particle $(N\bar{N})_{I=\frac{1}{2}} \leftrightarrow (\Lambda\bar{\Lambda})$ in low energy phenomena shows that the mass $(= 2m_N - [N\bar{N}]_{I=0} = 2m_\Lambda - [\Lambda\bar{\Lambda}])$ is much greater than the mass of pions. This is the reason why the mass formula is restricted to the particles in their "ground state" having various magnitudes of I and S except $I=0$ and $S=0$.

Table 2. The compositions, the masses and the stabilities for strong interactions of the particles having various I and S

Boson

| I | 0 | 1/2 | 1 | 3/2 |
|-----|---|--|--|---|
| S | 0 | 1/2 | 1 | 3/2 |
| 2 | ${}^+B_2^0 (\bar{A}\bar{A}pn)$ (1740) stable for $\rightarrow \theta^+ + \theta^0$ | | ${}^{++}B_2^1$ ${}^+B_2^1$ ${}^0B_2^1$ ($\bar{A}\bar{A}pp$) ($\bar{A}\bar{A}pn$) ($\bar{A}\bar{A}nn$) (1740) stable for $\rightarrow 2\theta$ | |
| 1 | | $\theta^+ (\bar{A}p)$ $\theta^0 (\bar{A}n)$ 965 (970) | | ${}^{++}B_1^{3/2}$ ${}^+B_1^{3/2}$ ${}^0B_1^{3/2}$ ${}^-B_1^{3/2}$ ($\bar{A}\bar{p}pp$) ($\bar{A}\bar{p}pn$) ($\bar{A}\bar{p}nn$) (1250) $\rightarrow \pi + \theta$ |
| 0 | — | | $\pi^+ (p\bar{n})$ $\pi^0 (p\bar{p} - n\bar{n})$ $\pi^- (p\bar{n})$ 273 $\left(\frac{p\bar{p} - n\bar{n}}{\sqrt{2}}\right)$ 273 (280) 264 (280) (280) | |
| -1 | | $\bar{\theta}^0 (\bar{A}\bar{n})$ $\bar{\theta}^+ (\bar{A}\bar{p})$ 965 (970) | | ${}^+B_{-1}^{3/2}$ ${}^0B_{-1}^{3/2}$ ${}^-B_{-1}^{3/2}$ ${}^{--}B_{-1}^{3/2}$ ($\bar{A}p\bar{n}\bar{n}$) ($\bar{A}p\bar{p}n$) ($\bar{A}p\bar{n}n$) (1250) $\rightarrow \pi + \bar{\theta}$ |
| -2 | ${}^-B_{-2}^0 (\bar{A}\bar{A}\bar{p}\bar{n})$ (1740) stable for $\rightarrow \bar{\theta}^+ + \bar{\theta}^0$ | | ${}^0B_{-2}^1$ ${}^-B_{-2}^1$ ${}^{--}B_{-2}^1$ ($\bar{A}\bar{A}\bar{n}\bar{n}$) ($\bar{A}\bar{A}\bar{p}\bar{n}$) ($\bar{A}\bar{A}\bar{p}\bar{p}$) (1740) stable for $\rightarrow 2\theta$ | |

Fermion

| I | 0 | 1/2 | 1 | 3/2 |
|-----|---|---|--|--|
| S | 0 | 1/2 | 1 | 3/2 |
| 2 | | ${}^{++}F_2^{1/2}$ ${}^+F_2^{1/2}$ ($\bar{A}\bar{A}ppn$) ($\bar{A}\bar{A}npn$) (4300) $\rightarrow N + 2\theta$ | | ${}^{++}F_2^{3/2}$ ${}^+F_2^{3/2}$ ${}^0F_2^{3/2}$ ${}^-F_2^{3/2}$ ($\bar{A}\bar{A}ppp$) ($\bar{A}\bar{A}ppn$) ($\bar{A}\bar{A}nnn$) (4300) $\rightarrow N + 2\theta$ |
| 1 | ${}^+F_1^0 (\bar{A}pn)$ (3160) $\rightarrow N + \theta$ | | ${}^{++}F_1^1$ ${}^+F_1^1$ ${}^0F_1^1$ ($\bar{A}pp$) ($\bar{A}pn$) ($\bar{A}nn$) (3160) $\rightarrow N + \theta$ | |
| 0 | | p 1840 n 1840 | | ${}^{++}F_0^{3/2}$ ${}^+F_0^{3/2}$ ${}^0F_0^{3/2}$ ${}^-F_0^{3/2}$ ($p\bar{p}\bar{n}$) ($p\bar{n}n$) 2400 (2120) $\rightarrow N + \pi$ |
| -1 | Λ 2180 | | $\Sigma^+ (\bar{A}\bar{n})$ $\Sigma^0 (\bar{A}\bar{p} - n\bar{n})$ $\Sigma^- (\bar{A}\bar{p}n)$ $\left(\frac{\bar{A}\bar{p} - n\bar{n}}{\sqrt{2}}\right)$ 2330 (2460) | |
| -2 | | $\Xi^0 (\bar{A}\bar{A}\bar{n})$ $\Xi^- (\bar{A}\bar{A}\bar{p})$ 2570 (2600) | | ${}^+F_{-2}^{3/2}$ ${}^0F_{-2}^{3/2}$ ${}^-F_{-2}^{3/2}$ ${}^{--}F_{-2}^{3/2}$ ($\bar{A}\bar{A}\bar{n}\bar{p}$) ($\bar{A}\bar{A}\bar{p}\bar{n}$) ($\bar{A}\bar{A}\bar{p}\bar{p}$) (2880) $\rightarrow \Xi + \pi$ |
| -3 | ${}^-F_{-3}^0 (\bar{A}\bar{A}\bar{A}\bar{p}\bar{n})$ (2800) stable for $\rightarrow \Xi + \bar{\theta}$ | | ${}^0F_{-3}^1$ ${}^-F_{-3}^1$ ${}^{--}F_{-3}^1$ ($\bar{A}\bar{A}\bar{A}\bar{n}\bar{n}$) ($\bar{A}\bar{A}\bar{A}\bar{p}\bar{n}$) ($\bar{A}\bar{A}\bar{A}\bar{p}\bar{p}$) (2800) stable for $\rightarrow \Xi + \bar{\theta}$ | |

B_b^a and F_b^a are the boson and the fermion having $I=a$ and $S=b$ respectively. The numerical values in the parenthesis are the mass values deduced from the mass formula (1) and the numerical values without the parenthesis are the observed mass values. The arrows show the disintegration processes through the strong interactions.

$$B_2^0, B_2^1 \rightarrow \theta + \pi, \tag{5}^*$$

$$B_{-2}^0, B_{-2}^1 \rightarrow \bar{\theta} + \pi, \tag{6}$$

$$F_{-3}^0, F_{-3}^1 \rightarrow \begin{cases} A + \bar{\theta} \\ \Sigma + \bar{\theta} \\ \Xi + \pi \end{cases}. \tag{7}^{**}$$

ii) On the $\theta-N$ scattering

There exist the particles F_1^0 and F_1^1 with mass of $3100 \sim 3200$. They disintegrate into $N + \theta$ through the strong interactions and so have very short lives. Their existence, however, gives rise to the resonance levels of $I=0$ and of $I=1$ in $\theta-N$ scattering in the region of the incident θ energy of about 300 Mev.***

For the $\theta^+ - p$ scattering, the cross section is given by the following,

$$\sigma^{\theta^+ + p \rightarrow \theta^+ + p}(\epsilon) = \pi \lambda^2 \cdot \frac{2 \times \text{spin of } F_1^1 + 1}{2 \times (2 \times \text{spin of } \theta^+ + 1)} \cdot \frac{I_{I=1}^2}{(\epsilon - \epsilon_r^{I=1})^2 + \frac{1}{4} I_{I=1}^2}, \tag{8}$$

where λ is the wave length of the incident θ , $\epsilon_r^{I=1}$ the resonance energy and $I_{I=1}$ the level width.

The role of F_1^0 and F_1^1 in $\theta-N$ scattering is analogous to that of $F_0^{3/2}$ in $\pi-N$ scattering. Furthermore, from the investigation of the excited levels of F_1^0 , F_1^1 , $F_0^{3/2}$ and $F_0^{1/2}$ (nucleon) we may have some further insight into the high energy reactions.

iii) On the spin of θ

Since in $\theta-N$ scattering there are two resonance levels of $I=0$ and of $I=1$ in the same energy regions, we cannot obtain a simple charge ratio of cross sections such as the one in $\pi-N$ scattering.

But the relation

$$\begin{aligned} & \frac{\sigma^{\theta^+ + p \rightarrow \theta^+ + p}(\epsilon_r^{I=1} \simeq 300 \text{ Mev})}{4\pi \lambda_r^2} \\ &= \frac{2 \times \text{spin of } F_1^1 + 1}{2 \times (2 \times \text{spin of } \theta^+ + 1)} (=W) \tag{9} \end{aligned}$$

Table 3. The value of W

| spin of θ \ spin of F_1^1 | 1/2 | 3/2 | 5/2 |
|------------------------------------|-----|-----|-----|
| 0 | 1 | 2 | 3 |
| 1 | 1/3 | 2/3 | 1 |
| 2 | 1/5 | 2/5 | 3/5 |

* (5) is consistent with the result obtained by Block: Bull. Am. Phys. Soc. 30 (1955), 13. Wang, Hsiao, Cheng, and Lü also discovered such a particle as type (6) with mass of about 1300: Acta Physica Sinica 11 (1955), 493.

** (7) is consistent with the result obtained by Eisenberg: Phys. Rev. 96 (1954), 541. If the masses of F_{-3}^0 and F_{-3}^1 are exactly 2800 as given by (1), then they do not decay into $\begin{cases} A + \bar{\theta} \\ \Sigma + \bar{\theta} \\ \Xi + \pi \end{cases}$. These particles have not yet been found. Bearing in mind the accuracy of (2) and supposing F_{-3}^0 and F_{-3}^1 are stable for the strong interactions, the masses of F_{-3}^0 and F_{-3}^1 must be found to be 3000~3500. The mass of Eisenberg's particle are 3000~3100.

*** It should also be remarked that the existence of such an isobar level with mass of about 4000 has already been found by the analysis of the energy spectrum of the K -meson produced by $N-N$ collision (Alichanow, private communication (1956) to Sakata).

gives us the method which restricts the spins of θ and F_1^1 , where λ_r is the wave length of the incident θ^+ with resonance energy $\epsilon_r^{I=1}$.

For the various spin values of θ and F_1^1 , W is given in Table 3.

In $\pi-N$ scattering, the relation

$$\frac{\sigma_{\pi^+ + p \rightarrow \pi^+ + p}(\epsilon_r = 190 \text{ Mev})}{4\pi\lambda_r^2} = \frac{2 \times \text{spin of } F_0^{3/2} + 1}{2 \times (2 \times \text{spin of } \pi^+ + 1)} = 2 \quad (10)$$

restricts uniquely the spin values of π and $F_0^{3/2}$ to 0 and $3/2$, respectively, as is easily seen from Table 3.

The author would like to express his sincere thanks to Prof. S. Sakata for his kind interest and valuable discussions.

References

- 1) T. Nakano and K. Nishijima, Prog. Theor. Phys. **10** (1953), 581.
K. Nishijima, Prog. Theor. Phys. **12** (1954), 107.
K. Nishijima, Prog. Theor. Phys. **13** (1955), 258.
M. Gell-Mann, Phys. Rev. **92** (1953), 883.
- 2) B. d'Espagnat and J. Prentki, Nuclear Phys. I (1956), 33.
B. d'Espagnat and J. Prentki, Phys. Rev. **99** (1955), 328.
- 3) M. Goldhaber, Phys. Rev. **101** (1956), 433.
- 4) M. A. Markov, Rep. Acad. Sci. U.S.S.R. (1955).
- 5) S. Sakata, Lecture at the annual meeting of the Phys. Soc. in Japan (1955).
- 6) S. Tanaka, Prog. Theor. Phys. **16** (1956), 625, 631.

On the Bose-Einstein Condensation

Shigetoshi KATSURA

Department of Applied Science, Tôhoku University, Sendai

(Received July 30, 1956)

We calculate the canonical partition function of the ideal Bose-Einstein gas by the modified method of steepest descent without the continuous spectrum approximation in the cases of periodic boundary condition and of perfectly reflecting wall condition. Above the critical point the steepest curve is the ordinary one through the saddle point. At and below the lambda temperature the steepest curve is such that the singularity of the integrand lies on it as a cusp, and the ordinary saddle point method loses its validity. The evaluation of the residual term in the integration, however, justifies the usual expression of the extensive quantities per particle and intensive quantities as N and V sufficiently large. In deriving the above results, a sufficient condition for the justification of the substitution of b_l^0 for $b_l^0(V)$ in canonical partition function of the imperfect gas is given and discussed.

§ 1. Introduction

The problem of the ideal Bose-Einstein condensation is treated by many authors since London¹⁾ tried to connect it to liquid helium. The theories of Bose condensation by grand canonical ensemble without the continuous spectrum approximation were presented by Fowler-Jones,²⁾ and by de Groot-Hooyman-Seldam.³⁾ De Groot et al. state that their method is not justified because of the use of the grand canonical distribution. We think with Landsberg,⁴⁾ that in these methods, N , which they think as given, is appropriate to be considered as the mean number \bar{N} for the given fugacity z . In the theories using the method of canonical ensemble one can count the works by Frazer⁵⁾ and by Schubert.⁶⁾ Frazer refines the method of Kahn and Uhlenbeck. Schubert points out that the saddle point method loses its applicability below the critical temperature. Dingle⁷⁾ tempts the refinement of the saddle point method. There seems, however, to be some room for discussions.

The author treats here the system of ideal Bose particles taking the comparison to imperfect gas into consideration. The purpose is the derivation of the thermodynamic properties, especially the mathematical grasp of the condensation.

First a sufficient condition for the substitution of b_l^0 for $b_l(V)$ —the generalized cluster coefficients—in canonical partition function is given discussing the theory of condensation of ideal Bose gas and of imperfect gas. Then the modified method of steepest descent is applied. At and below the lambda temperature the ordinary saddle point method loses its validity and hence the contour is taken so that it indents a singularity of integrand. The result justifies the usual expression of the thermodynamic quantities per particle.

§ 2. A remark on the volume dependence of the cluster integrals

We consider the system composed of N particles in the container whose volume is V . The content of this section does not concern whether they have individuality (classical statistics) or not (quantum statistics), and whether they have mutual interactions or not. Assigning the energy eigenvalue of the whole system E_l , we have the canonical partition function (abbreviated as C.P.F. hereafter) at temperature T ,

$$\mathcal{Q}^{\dagger}(T, V, N) = \sum e^{-E_k/kT}. \quad (2.1)$$

The grand partition function (abbreviated as G.P.F. hereafter) of the system at the fugacity z is,

$$\Xi^{\dagger}(T, V, z) = \sum_{N=0}^{\infty} \mathcal{Q}^{\dagger}(T, V, N) z^N. \quad (2.2)$$

We define the coefficient of z^l in the expansion of $1/V \cdot \log \Xi(T, V, z)$ as the generalized cluster coefficient, and denote it by $b_l^{\dagger}(T, V)$.

$$\Xi^{\dagger}(T, V, z) = \exp \left[V \sum_{l=1}^{\infty} b_l^{\dagger}(V, T) z^l \right]. \quad (2.3)$$

In the case of interacting classical particles $b_l(T, V)$ is identified with the cluster integral in which the volume dependence is taken into consideration.

Then $\mathcal{Q}^{\dagger}(T, V, N)$ is given by

$$\mathcal{Q}^{\dagger}(T, V, N) = \sum \left(\sum_{l=1}^N l m_l = N \right) \prod_{l=1}^N \frac{(V b_l^{\dagger}(V, T))^{m_l}}{m_l!}. \quad (2.4)$$

Let

$$\lim_{V \rightarrow \infty} b_l^{\dagger}(V, T) = b_l^0(T). \quad (2.5)$$

The expressions in which b_l^0 is used instead of $b_l(V)$ in (2.3) and (2.4) are denoted by $\Xi^{(0)}(T, V, z)$ and $\mathcal{Q}^{(0)}(T, V, N)$ respectively.

$$\Xi^{(0)}(T, V, z) = \exp \left[V \sum_{l=1}^{\infty} b_l^0(T) z^l \right], \quad (2.6)$$

$$\mathcal{Q}^{(0)}(T, V, N) = \sum \left(\sum_{l=1}^N l m_l = N \right) \prod_{l=1}^N \frac{(V b_l^0(T))^{m_l}}{m_l!}. \quad (2.7)$$

Usually (2.6) and (2.7) are treated in place of (2.4) for the sake of simplicity. Generally, the results from (2.3), (2.4), (2.6) and (2.7) are not necessarily the same⁽⁸⁾⁽⁹⁾⁽¹⁰⁾ even for infinitely large system. Here we give a sufficient condition for justifying the substitution of b_l^0 for $b_l(V)$.*

We fix the temperature T (which is assumed to exist) such that $b_l^0(T)$ is positive for all l . If positive constants a and δ can be found as:

* The author is indebted to Ikeda's paper.⁽¹⁰⁾

$$\exp(-aIV^{-\delta}) < b_i^\dagger(V)/b_i^0 < \exp(aIV^{-\delta}), \quad (2.8)$$

then

$$|\log \mathcal{Q}^\dagger(T, Nv, N) - \log \mathcal{Q}^{(0)}(T, Nv, N)| = O(N^{\text{Max}(1/2, 1-\delta)}), \quad (2.9)$$

that is,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{Q}^\dagger(T, Nv, N) = \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathcal{Q}^{(0)}(T, Nv, N). \quad (2.9')$$

Proof: We write

$$\begin{aligned} \prod_{l=1}^N \frac{(Vb_l^\dagger(V))^{m_l}}{m_l!} &= T^\dagger\{m_l\}, \\ \prod_{l=1}^N \frac{(Vb_l^0)^{m_l}}{m_l!} &= T^{(0)}\{m_l\}. \end{aligned} \quad (2.10)$$

A set of the values of m_l which makes $T^\dagger\{m_l\}$ a maximum (rather the greatest) subject to $\sum_{l=1}^N lm_l = N$, is denoted by $\{m_l'\}$, and that corresponding to $T^{(0)}\{m_l\}$ by $\{m_l''\}$.*

Then

$$T^\dagger\{m_l'\} \geq T^\dagger\{m_l''\}, \quad T^{(0)}\{m_l''\} \geq T^{(0)}\{m_l'\} \quad (2.10')$$

$$\begin{aligned} &|\log \mathcal{Q}^\dagger(T, Nv, N) - \log \mathcal{Q}^{(0)}(T, Nv, N)| \\ &\leq |\log \sum T^\dagger\{m_l\} - \log T^\dagger\{m_l'\}| \\ &\quad + |\log \sum T^{(0)}\{m_l\} - \log T^{(0)}\{m_l''\}| \\ &\quad + |\log T^\dagger\{m_l'\} - \log T^{(0)}\{m_l''\}|. \end{aligned} \quad (2.11)$$

Since

$$T^\dagger\{m_l'\} < \sum T^\dagger\{m_l\} < P_N T^\dagger\{m_l'\},$$

where P_N is the partition number of N ,

$$|\log \sum T^\dagger\{m_l\} - \log T^\dagger\{m_l'\}| < \log P_N \sim \pi(2N/3)^{1/2}. \quad (2.12)$$

On the other hand, by (2.8)

$$\begin{aligned} |\log T^\dagger\{m_l'\} - \log T^{(0)}\{m_l''\}| &= \sum m_l' \left| \log \frac{b_l^\dagger(Nv)}{b_l^0} \right| \\ &\leq aN^{1-\delta} v^{-\delta}. \end{aligned} \quad (2.13)$$

From (2.10'), (2.11), (2.12), (2.13) and corresponding relations for $\{m_l''\}$, (2.9) and (2.9') are established.

* One of the sets of m_l is denoted by $\{m_l'\}$, when there exist more than two sets which make $T^\dagger\{m_l\}$ maxima.

Corollary. Even when positive constants a and δ satisfying (2.8) cannot be found for all positive V and l , there results (2.9) for $v > v_0$, provided that positive constants a , δ and v_0 can be found in such a manner that (2.8) is satisfied for $V > lv_0$. (cf. § 8 and Fig. 3).

§ 3. Cluster coefficient of ideal Bose gas

We consider the system of non-interacting Bose particles contained in the cube whose each edge length is $L(=V^{1/3})$. The periodic boundary condition and the condition of perfectly reflecting wall are considered in parallel. The upper indices P and R denote the former and the latter case respectively. When the boundary condition is not concerned but only the volume dependence is taken into account, the upper index 1 is attached. In this section $b_l(T, V)$ of Bose particles is obtained and (2.9) is proved for \mathcal{Q}^P and for \mathcal{Q}^R by the theorem in the preceding section and by a slightly modified method respectively.

The energy eigenvalues in the state (l, m, n) of one particle problem are given by

$$\epsilon_{lmn}^P = h^2(l^2 + m^2 + n^2)/2mL^2 \quad (3.1)$$

for the periodic boundary conditions, and

$$\epsilon_{lmn}^R = h^2(l^2 + m^2 + n^2)/8mL^2 \quad (3.1')$$

for the perfectly reflecting wall conditions respectively, where l, m, n are positive and negative integers including zero for the former case and positive integers for the latter case. When we denote the number of particles in the state (l, m, n) by n_{lmn} , C. P. F. and G. P. F. are given by

$$\mathcal{Q}^1(T, V, N) = \sum (\sum n_{lmn} = N) \exp \left[- \sum_{lmn} n_{lmn} \epsilon_{lmn}/kT \right] \quad (3.2)$$

and

$$\Xi^P(T, V, z) = 1 / \prod_{-\infty}^{\infty} \prod_{-\infty}^{\infty} \prod_{-\infty}^{\infty} [1 - z \exp \{ -\pi(l^2 + m^2 + n^2)/s \}], \quad (3.3)$$

$$\Xi^R(T, V, z) = 1 / \prod_1^{\infty} \prod_1^{\infty} \prod_1^{\infty} [1 - z \exp \{ -\pi(l^2 + m^2 + n^2)/4s \}], \quad (3.3')$$

respectively, where

$$s = (L/\lambda)^2, \quad \lambda = h/(2\pi mkT)^{1/2}. \quad (3.4)$$

Since $\sum_l \sum_m \sum_n \exp \{ -\pi(l^2 + m^2 + n^2)/s \}$ converges for all finite s , the infinite products in the denominator of (3.3) converge in all domain of complex z plane, and have zeros at

$$z = \exp \{ \pi(l^2 + m^2 + n^2)/s \}. \quad (3.5)$$

Therefore the analytic continuation of $\Xi^P(T, V, z)$ exists in all complex z plane except at poles given by (3.5). The similar can be said for $\Xi^R(T, V, z)$.

From (3.3) and (3.3') the generalized cluster coefficients $b_l(V)$ are expressed by Theta function.²⁾ In the case of periodic conditions

$$b_l^P(T, V) = \left[\sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} \exp \{ -l\pi(l^2 + m^2 + n^2)/s \} \right] / lV \\ = [\vartheta_3(0, il/s)]^3 / lV. \quad (3.6)$$

It can be expressed by the inverse transformation of Theta function

$$= [(s/l)^{1/2} \vartheta_3(0, is/l)]^3 / lV. \quad (3.7)$$

In the case of perfectly reflecting wall conditions,

$$b_l^R(T, V) = [\vartheta_3(0, il/4s) - 1]^3 / 8lV \quad (3.8)$$

$$= [(4s/l)^{1/2} \vartheta_3(0, i4s/l) - 1]^3 / 8lV. \quad (3.9)$$

In the limit $V \rightarrow \infty$,

$$\lim_{V \rightarrow \infty} b_l^P(T, V) = \lim_{V \rightarrow \infty} b_l^R(T, V) = b_l^0(T) = \lambda^{-3} l^{-5/2}. \quad (3.10)$$

$b_l^P(T, V)$ and $b_l^R(T, V)$ are positive for all T and V . The former tends to $b_l^0(T)$ decreasingly and the latter tends to it increasingly.

The comparison of the series expansion of (3.6) and (3.7) to $\epsilon^{l/s}$ yields

$$1 < b_l^P(V) / b_l^0 < \exp(3 \lambda^2 l V^{-2/3}) \quad (3.11)$$

for $0 < s/l < \infty$. Thus (2.9) is proved for the ideal Bose gas in the case of periodic conditions.

On the other hand, one cannot find constants a and δ satisfying (2.8) through the whole $V-l(>0)$ plane in the case of perfectly reflecting wall conditions.

For $s/l < 1$,

$$\vartheta_3(0, i4s/l) - (l/4s)^{1/2} > \exp(-\pi l/4s). \quad (3.12)$$

For $1 < s/l < \infty$,

$$\vartheta_3(0, i4s/l) - (l/4s)^{1/2} > \exp(-\pi(l/4s)^{1/2}). \quad (3.13)$$

When N and V are large enough, s is smaller than N .

$$|\log T^R \{m_l'\} - \log T^{(0)} \{m_l'\}| \\ = \sum_{l=1}^N m_l' |\log(b_l^R(V) / b_l^{(0)})| \\ \leq \sum_{l=1}^{l \leq s} l m_l' (3\pi/4s) + \sum_{l \geq s}^N \sqrt{l} m_l' (3\pi/(4s)^{1/2}) \\ \leq 3\pi N \left(\frac{1}{4s} + \frac{1}{\sqrt{4s}} \right). \quad (3.14)$$

Substitution of (2.12) and (3.14) in (2.11) yields

$$\left| \frac{1}{N} \log \Omega^R(T, Nv, N) - \frac{1}{N} \log \Omega^{(0)}(T, Nv, N) \right| < O(N^{-1/3}). \quad (3.15)$$

Relations of (3.11), (3.12) and (3.13) are shown in Fig. 1.

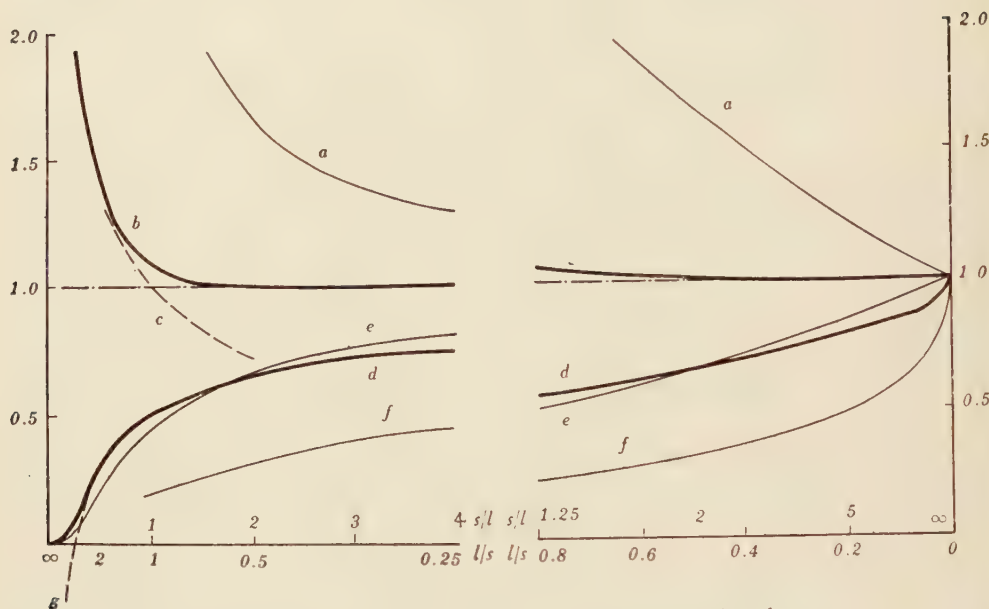


Fig. 1a

- a. $e^{l/s}$
- b. $\vartheta_3(0, is/l)$
- c. $(l/s)^{1/2}$
- d. $\vartheta_3(0, il/4s) - (l/4s)^{1/2}$

Fig. 1b

- e. $\exp[-\pi l/4s]$
- f. $\exp[-\pi(l/4s)^{1/2}]$
- g. $1 - (l/4s)^{1/2}$

§ 4. Determination of the contour

In this section we evaluate $\Omega^{(0)}(T, V, N)$ by the modified method of steepest descent. We denote the analytic function defined by a power series $\sum_{l=1}^{\infty} z^l/l^s$ by $\phi(z, s)$. The properties of this function are discussed by Lindelöf,⁽¹¹⁾ Truesdell⁽¹²⁾ and others.⁽¹³⁾⁽¹⁴⁾⁽¹⁵⁾ When C is the contour enclosing $z=0$ and such that $|z| \sim 1-\delta$ on it, we have from (2.6) and (2.7)

$$\Omega^{(0)}(T, V, N) = -\frac{1}{2\pi i} \int_C e^{\Delta f(z)} dz, \quad (4.1)$$

where

$$f(z) = \phi(z, 5/2) - \rho \log z, \quad (4.2)$$

$$A = V\lambda^{-3}, \quad \rho = \lambda^3(N+1)/V = \lambda^3/v. \quad (4.3)$$

Now the contour is transformed so that the imaginary part of $f(z)$ is zero and the real part of $f(z)$ tends to $-\infty$ on both ends of the contour. Write $z = re^{i\theta}$, $\log z = R e^{i\theta}$, where r, θ, R and θ are real, and the equation of the contour is,

$$\operatorname{Im} f(z) = \operatorname{Im} \phi(re^{i\theta}, 5/2) - \rho\theta = 0. \quad (4.4)$$

Lindelöf expansion of $\phi(z, s)^{(11)(12)(14)}$ is well known,

$$\phi(z, s) = \Gamma'(1-s) (-\log z)^{s-1} + \sum_{n=0}^{\infty} \zeta(s-n) \frac{(\log z)^n}{n!}. \quad (4.5)$$

Then

$$f(z) = \zeta(5/2) + (\rho_s - \rho) R e^{i\theta} + 4\sqrt{\pi}/3 \cdot R^{3/2} e^{i\zeta(3/2)(\pi + \theta)/2} \\ - 0.730 R^2 e^{i2\theta} - 0.0347 R^3 e^{i3\theta} + O(R^4), \quad \rho_s = \zeta(3/2). \quad (4.6)$$

The shape of the contour and $\operatorname{Re} f(z)$ are symmetrical with regard to the real axis.

When r is large ($\log r \gg \pi$), (4.4) is transformed to

$$\pi - \theta = 3\sqrt{\pi}/4 \cdot \rho (\log r)^{-3/2} + O((\log r)^{-2}) \quad (4.7)$$

and there,

$$\operatorname{Re} f(z) = -8/15 \sqrt{\pi} \cdot (\log r)^{5/2} - \rho \log r + O((\log r)^{-1/2}). \quad (4.8)$$

(4.7) and (4.8) are established by the asymptotic expansion of $\phi(z, s)$ given in Appendix.

When we investigate the contour near the real axis, we have three distinct cases to consider, in which ρ is less than, equal to, or greater than $\zeta(3/2) \equiv \rho_s$, respectively.

1° When $\rho < \zeta(3/2) = \rho_s$ (above the lambda temperature), (4.1) is calculated by the ordinary saddle point method. The saddle point is given by z_0 at which

$$z_0 f'(z_0) = \phi(z_0, 3/2) - \rho = 0. \quad (4.9)$$

The contour is orthogonal to the real axis at $z = z_0$, since $f''(z_0) = \phi(z_0, \frac{1}{2})/z_0^2 > 0$.

2° The case $\rho = \rho_s$: The second term in the right-hand side of (4.6) vanishes. The steepest curve on which $\operatorname{Re} f(z)$ takes the greatest value at $z=1$ is given by

$$\theta = 7\pi/3 + 0.730/2 \sqrt{\pi} \cdot R^{1/2} + O(R) \quad (4.10)$$

when R is small. On this curve, by (4.5) and (4.10), we have

$$\operatorname{Re} f(z) = \zeta(5/2) - 4\sqrt{\pi}/3 \cdot R^{3/2} + 0.730/2 \cdot R^2 + O(R^3). \quad (4.11)$$

3° The case $\rho > \rho_s$: The steepest curve on which $\operatorname{Re} f(z)$ takes the greatest value at $z=1$ is given by

$$\theta = 2\pi + 4\sqrt{\pi}/3 (\rho - \rho_s) \cdot R^{1/2} + O(R), \quad (4.12)$$

when R is small. On this curve, by (4.5) and (4.12),

$$\operatorname{Re} f(z) = \zeta(5/2) - (\rho - \rho_s) R - 4\sqrt{\pi}/3 \cdot R^{3/2} + O(R^2). \quad (4.13)$$

The shape of the contour is shown in Fig. 2. When $\rho < \rho_s$, it is orthogonal to the real axis at z_0 ($0 < z_0 < 1$), where $\operatorname{Re} f(z)$ takes its maximum value. When $\rho = \rho_s$,

it has a singularity of the integrand at the cusp $z=1$ and crosses the real axis at $\pm\pi/3$. When $\rho > \rho_s$, it has a real axis as a tangent line at $z=1$. In the latter two cases, $\operatorname{Re} f(z)$ takes at $z=1$ the greatest value on the contour. In all cases the contour goes to infinity ($\operatorname{Re} z \rightarrow -\infty$) by (4.5).

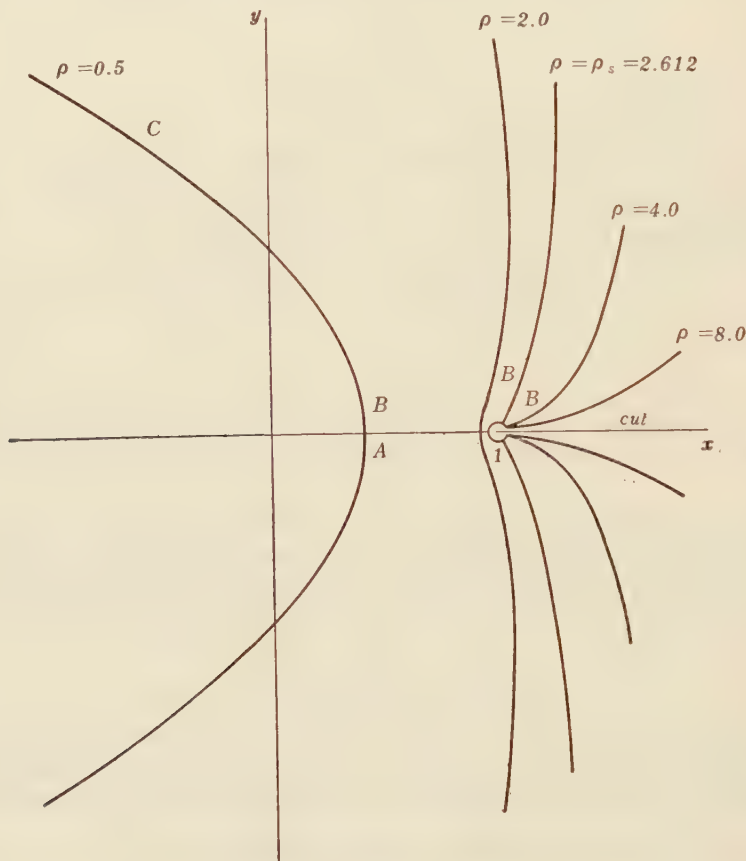


Fig. 2. Contour in z plane

§ 5. Integration by the modified steepest descent method

Now we consider the integration on the contour given in the preceding section. When $\rho < \rho_s$, it is an ordinary saddle point method. When $\rho \geq \rho_s$, the ordinary saddle point loses its validity, and the path of integration is taken along the curve indented by a small circle at $z=1$. The integration on this small circle, however, tends to zero as the radius of this circle tends to zero. Therefore we can integrate it over the curve starting from $z=1$. Such an integration can be shown to be also of order of $\exp[Af(1) + a \log .1]$ by taking a proper choice of ε , the main contributing region to the integrand.

We denote z_0 ($\rho < \rho_s$) or 1 ($\rho \geq \rho_s$) by A . The points whose arc distances from A

on the steepest curve are ε , L , and ∞ , are denoted by B , C and D respectively. The arc length from A along the steepest curve is denoted by t . c_i and M_i are appropriate constants hereafter.

1° The case $\rho < \rho_s$.

On AB

$$f(z_0) - c_1 t^2 + M_1 t^3 \geq \operatorname{Re} f(z) \geq f(z_0) - c_1 t^2 - M_1 t^3. \quad (5.2)$$

On BC

$$\operatorname{Re} f(z) \geq f(z_0) - c_1 t^2 - M_2 t^3, \quad (5.3)$$

where $c_1 = \frac{1}{2} f''(z_0)$.

Let $\varepsilon = A^{-2/3}$, then

$$\begin{aligned} \int_{AB} e^{\Lambda f(z)} dz &= e^{\Lambda f(z_0)} i(\pi/\Lambda c_1)^{1/2} (1 + O(A^{-1/5})) \Phi(\varepsilon \sqrt{\Lambda c_1}) \\ &= e^{\Lambda f(z_0)} i(\pi/\Lambda c_1)^{1/2} (1 + O(A^{-1/5})) \{1 + O(\exp(-c_1 A^{1/5}))\}, \end{aligned} \quad (5.4)$$

$$\int_{BC} e^{\Lambda f(z)} dz = e^{\Lambda f(z_0)} O(\exp(-c_1 A^{1/5})), \quad (5.5)$$

where $\Phi(x)$ is the error function.

2° The case $\rho = \rho_s$.

From (4.10) and (4.11) we have on AB

$$dz = \{e^{i\pi/3} + O(R^{1/2})\} dR, \quad (5.6)$$

$$f(z_0) - c_2 R^{3/2} + M_3 R^2 \geq \operatorname{Re} f(z) \geq f(z_0) - c_2 R^{3/2} - M_3 R^2, \quad (5.7)$$

and on BC

$$\operatorname{Re} f(z) \geq f(z_0) - c_2 t^{3/2} - M_4 t^2. \quad (5.8)$$

Let $\varepsilon = A^{-7/12}$, then

$$\begin{aligned} \int_{AB} e^{\Lambda f(z)} dz &= e^{\Lambda f(1)} \int_0^\varepsilon e^{-\Lambda c_2 R^{3/2}} dR \cdot e^{i\pi/3} (1 + O(A^{-1/6})) \\ &= e^{\Lambda f(1)} e^{i\pi/3} \cdot 2/3 \cdot (\Lambda c_2)^{-2/3} \gamma(2/3, \varepsilon_2 \Lambda \varepsilon^{3/2}) (1 + O(A^{-1/6})) \end{aligned} \quad (5.9)$$

$$\begin{aligned} &= e^{\Lambda f(1)} e^{i\pi/3} \cdot \sqrt{\pi}/3 \cdot (c_2 A^{1/8})^{-2/3} \\ &\quad \times (1 + O(A^{-1/6})) \{1 + O(\exp(-c_2 A^{1/8}))\}. \end{aligned} \quad (5.10)$$

where $\gamma(n, z)$ is the incomplete gamma function.

3° The case $\rho > \rho_s$.

From (4.12) and (4.13) we have on AB

$$dz = (1 + O(R^{1/2})) dR, \quad (5.12)$$

$$f(z_0) - c_3 R + M_4 R^{3/2} \geq \operatorname{Re} f(z) \geq f(z_0) - c_3 R - M_4 R^{3/2},$$

where $c_3 = \rho - \rho_s$. And on BC

$$\operatorname{Re} f(z) \geq f(z_0) - c_3 t - M_5 t^{3/2}. \quad (5.13)$$

Let $\varepsilon = A^{-4/5}$, we have

$$\begin{aligned} \int_{AB} e^{\Lambda f(z)} dz &= e^{\Lambda f(1)} \int_0^\varepsilon e^{-\Lambda c_3 R} dR (1 + O(A^{-1/5})) \\ &= e^{\Lambda f(1)} \cdot 1/\Lambda c_3 \cdot (1 + O(A^{-1/5})) (1 - \exp(-c_3 A^{1/5})), \end{aligned} \quad (5.14)$$

$$\int_{BC} e^{\Lambda f(z)} dz = e^{\Lambda f(1)} O(\exp(-c_3 A^{1/5})). \quad (5.15)$$

Integration on CD . Irrespective of $\rho \equiv \rho_s$, we have from (4.13) and (4.14)

$$dz = -\exp\{3\sqrt{\pi}\rho/4 \cdot (\log r)^{-3/2}\} \{1 + O((\log r)^{-5/2})\} dr, \quad (5.16)$$

$$\begin{aligned} \int_{CD} e^{\Lambda f(z)} dz &= \int_L^\infty \exp[-\Lambda c_4 (\log r)^{5/2} + O((\log r)^{-1/2})] \\ &\quad \times \{1 + O((\log r)^{-5/2})\} dr \\ &< \int_L^\infty e^{-c_4 \Lambda \log^2 r} dr = O(\exp(-c_5 A)). \quad (c_5 = c_4 \log L) \end{aligned} \quad (5.17)$$

Thus we get the following results:

For $\rho < \rho_s$ (above the lambda temperature),

$$\mathcal{Q}^{(0)}(T, V, N) = A_1 A^{-1/2} \exp[V\lambda^{-3}\phi(z_0, 5/2) - (N+1)\log z_0] \cdot (1 + O(A^{-1/5})), \quad (5.18)$$

where z_0 is given by (4.9).

For $\rho = \rho_s$ (at the lambda temperature),

$$\mathcal{Q}^{(0)}(T, V, N) = A_2 A^{-3/4} \exp[V\lambda^{-3}\zeta(5/2)] \cdot (1 + O(A^{-1/5})). \quad (5.19)$$

For $\rho > \rho_s$ (below the lambda temperature),

$$\mathcal{Q}^{(0)}(T, V, N) = A_3 A^{-1} \exp[V\lambda^{-3}\zeta(5/2)] \cdot (1 + O(A^{-1/5})). \quad (5.20)$$

Where A_1 , A_2 and A_3 are constants. From (5.18), (5.19), (5.20) and (2.9), (3.15) we have

$$\log \mathcal{Q}^{(0)}(T, V, N) = V\lambda^{-3}\phi(z_0, 5/2) - N \log z_0 + O(\log A) \quad (5.21)$$

where z_0 is given by (4.9) for $\rho < \rho_s$ and $z_0 = 1$ for $\rho \geq \rho_s$.

When we make N and V large for fixed v and T , $1/N \cdot O(\log A)$ become sufficiently small. This state of affairs is the same for the relations obtained from the differentiation of (5.21) by z or T . Thus we have proved by the modified steepest descent method that the extensive thermodynamic quantities per particle and intensive thermodynamic quantities are given by the usual expressions at the limit N and $V \rightarrow \infty$ under constant $v = V/N$.

§ 6. Relation to Yang-Lee's theory

In this section we give a remark on the method of the grand partition function. The analytic continuation of G. P. F. given by (3.3) is a one-valued function of complex

z plane except at the poles which lie on the part $z > 1$ of the real axis given by (3.5), that is, the condensation point is the end point of the distribution function of these poles. It is an interesting contrast to the fact that the condensation point is the cross point of the distribution of zeros and the real axis in the case of lattice gas.^{16)*}

§ 7. Virial expansion

From (5.21), isotherm of the gaseous phase is given by

$$\gamma \equiv \frac{p}{kT} \frac{1}{\lambda^3} = \phi(z, 5/2), \quad \rho \equiv \frac{1}{v} \frac{1}{\lambda^3} \phi(z, 3/2). \quad (7.1)$$

Elimination of z from (7.1) gives the virial expansion:

$$\frac{p}{kT} = \frac{1}{v} \left[1 - \sum_{k=1}^{\infty} \frac{k}{k+1} \beta_k \frac{1}{v^k} \right]. \quad (7.2)$$

The explicit expression of (7.2) is given by London¹⁾ and by Widom.^{17)**}

The shape of the curve given by (7.2) near $\rho \simeq \rho_s$ can be obtained by eliminating $\log z$ from Lindelöf expansion of (7.1):

$$\begin{aligned} \gamma - \zeta(5/2) &= -\zeta(3/2)/4\pi \cdot (\rho - \rho_s)^2 \\ &\quad - (1/6\pi + \zeta(1/2)\zeta(3/2)/24\pi^2) (\rho - \rho_s)^3 + \cdots \end{aligned} \quad (7.3)$$

For $\rho < \rho_s$ (7.3) gives the mapping between the first leaves of the Riemannian surfaces of (7.1), and for $\rho > \rho_s$ it gives that between the second leaves of (7.1). From (7.3)

$$(d\gamma/d\rho)_{\rho=\rho_s} = 0. \quad (7.4)$$

That is, not the singularity of $\sum_{k=1}^{\infty} \beta_k \rho^k$, but the point which satisfies $\sum_{k=1}^{\infty} k \beta_k \rho^k = 1$, corresponds to the singularity of $\sum_{l=1}^{\infty} b_l z^l$. (7.3) is continued smoothly from the region $\rho < \rho_s$ to the region $\rho > \rho_s$. The trace of (7.1) on the second leaves of Riemannian surface shows that the analytic continuation of $\gamma = \gamma(\rho)$ has no singularity on positive real axis and $\gamma(\infty) = -\infty$. The complex value of ρ which corresponds to $z=1$ on the third leaf of z

* In this point the lattice gas is rather similar to Husimi-Temperley model⁹⁾ than to the ideal Bose gas.

** β_k is related to b_l by

$$\beta_k = \frac{(-)^{k+1}}{k} \frac{1}{k!} \begin{vmatrix} k \cdot 2b_2 & 1 & 0 & 0 & 0 & \cdots \\ 2k \cdot 3b_3 & (k+1) \cdot 2b_2 & 2 & 0 & 0 & \cdots \\ 3k \cdot 4b_4 & (2k+1) \cdot 3b_3 & (k+2) \cdot 2b_2 & 3 & 0 & \cdots \\ 4k \cdot 5b_5 & (3k+1) \cdot 4b_4 & (2k+2) \cdot 3b_3 & (k+3) \cdot 2b_2 & 4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{vmatrix}$$

This is a generalization of Widom's results. It is valid irrespective of the explicit form of b_l .

(which is connected at the logarithmic singularity $z=0$ to the second leaf) is a singularity of (7.2).*

§ 8. Discussions of the theory of condensation of the classical imperfect gas

The present analysis of the behavior of the ideal Bose gas reveals some points in question in the theory of condensation of the classical imperfect gas.

Ikeda¹⁰⁾ derives a sufficient condition for justifying a substitution of b_l^0 for $b_l(V)$ in the canonical partition function. His theorem (corollary 2 in reference¹⁰⁾) states that the normal property such as

$$0 \leq b_l(V) \leq b_l^0 \quad (8.1)$$

for

$$V \geq lv_s \quad (8.2)$$

where $v_s = 1/\sum_{l=1}^{\infty} l b_l z_s^l$ and z_s is the smallest real positive singularity of the analytic function defined by $\sum_{l=1}^{\infty} b_l z^l$, is a sufficient condition for (2.9) in $v > v_s$.

This theorem should be interpreted rather as: When we can take v_0 independent of l in such a way that the normal property (8.1) is satisfied for $V \geq lv_0$, then (2.9) holds for $v > v_0$. Here we may take as a normal property, for example, Ikeda's condition, Frazer's condition,⁵⁾ or condition (2.8) of the present paper, etc. (Cf. Fig. 3). It is the

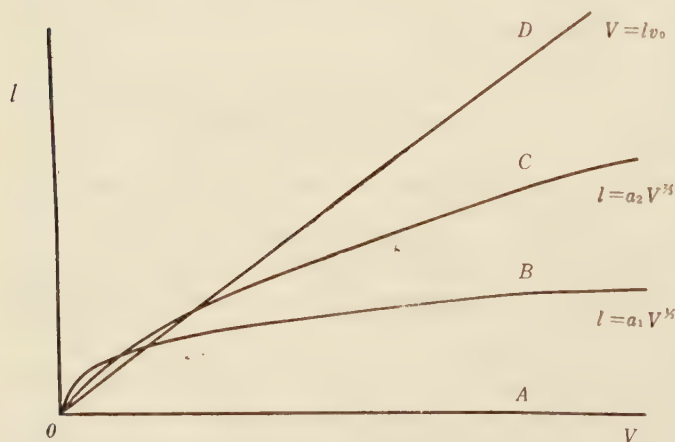


Fig. 3. Ikeda proves the normal property of $b_l(V)$ in the domain AOB. When $b_l(V)$ is normal also in BOD, (2.9) is true for $v > v_0$.

problem in future to confirm whether v_0 is greater than v_s , and whether there exists such v_0 . (For example, such v_0 does not exist when the border line between the normal and abnormal region of $b_l(V)$ is given by $l \sim V^{2/3}$).

It should further be noticed that the convergence of $b_l(V) \rightarrow b_l^0$ is not a sufficient condition for (2.9), and that (8.1) with (8.2) is not a sufficient condition for the coincidence between the results by

C. P. F. and by G. P. F. Such examples for $b_l(V)$ are easily constructed.

* Shapiro and Fuchs proves the finiteness and $12.56 \dots < R < 27.73 \dots$ of the radius of the convergence R of (7.1).¹⁸⁾

§ 9. Conclusion

Firstly, a sufficient condition for the justification of the replacement of $\mathcal{Q}^{(0)}(T, V, N)$ for $\mathcal{Q}^1(T, V, N)$ has been given. Next $\mathcal{Q}^{(0)}(T, V, N)$ of the ideal Bose gas, for which the above replacement is permitted, has been obtained by the modified method of the steepest descent. Above the lambda temperature the path of the integration is an ordinary steepest curve crossing the saddle point, while at and below the lambda temperature it has a singularity of the integrand as a cusp, and the ordinary saddle point method loses its validity. The usual expressions of extensive thermodynamic quantities per particle and intensive thermodynamic quantities, however have been proved to be true in the limit of $N \rightarrow \infty$ under constant V/N by the present method. It has been shown also that the density at the beginning of the condensation is not the singularity of the virial series, and that the value of the fugacity at the beginning of the condensation is the end point of the distribution function of the poles of the grand partition function. The theory of the imperfect gas has been briefly discussed.

It is the author's pleasure to thank Prof. Y. Nomura for his encouragement and many discussions of this problem with the author.

Appendix. Asymptotic expansion of $\phi(z, s)$

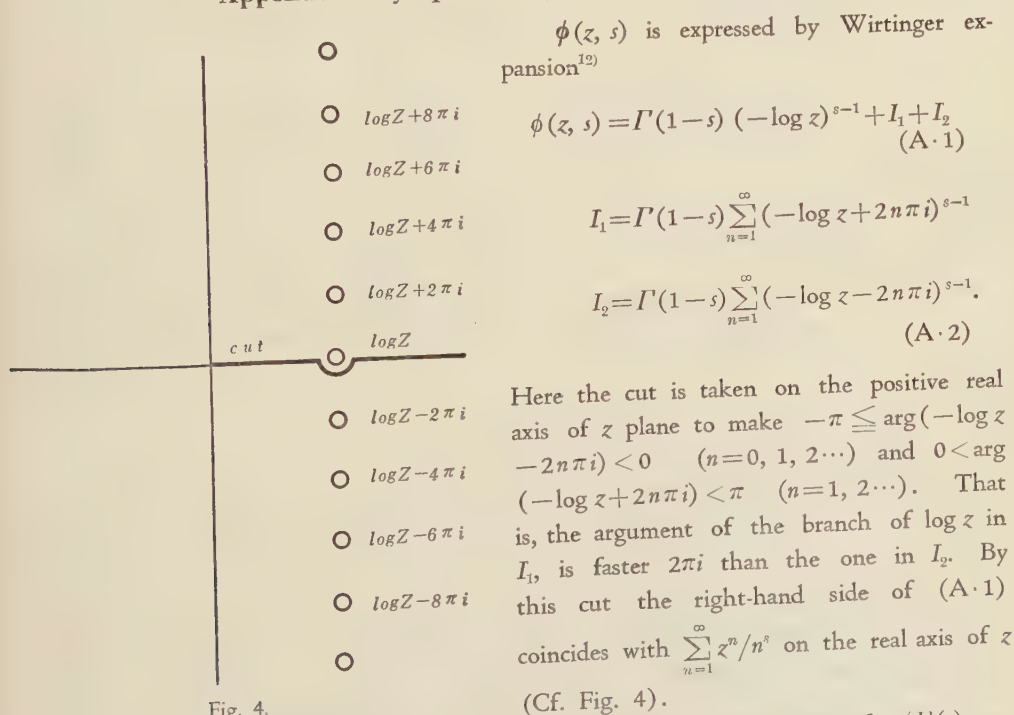


Fig. 4.

Changing the integration variable t to $t\tilde{z}$ in Hankel representation of $1/\Gamma'(s)$, we

have

$$(\hat{\zeta} e^{-\pi i})^{s-1} = \frac{\Gamma(s)}{2\pi i} \int_{\infty \exp i\delta}^{(0+)} t^{-s} e^{-t\hat{\zeta}} dt, \quad (\text{A} \cdot 3)$$

where $-\pi/2 - \arg \hat{\zeta} < \delta < \pi/2 - \arg \hat{\zeta}$, $\delta \leq \arg t \leq 2\pi + \delta$.

By (A.3), I_2 is transformed to

$$I_2 = \frac{1}{2i \sin \pi s} \int_{\infty \exp i\delta_2}^{(0+)} t^{-s} e^{-t \log z} \sum_{n=1}^{\infty} e^{-2\pi i n t} dt. \quad (\text{A} \cdot 4)$$

Since $\sum_{n=1}^{\infty}$ in the integrand converges when $t \rightarrow \infty \exp i\delta_2$,

$$= \frac{1}{2i \sin \pi s} \int_{\infty \exp i\delta_2}^{(0+)} t^{-s} e^{-t \log z} \frac{1}{e^{2\pi i t} - 1} dt. \quad (\text{A} \cdot 5)$$

The integrand is expanded in Bernoulli numbers,*

$$= \frac{1}{2i \sin \pi s} \int_{\infty \exp i\delta_2}^{(0+)} t^{-s} e^{-t \log z} \times \left[\frac{1}{t} - \frac{1}{2} - \sum_{n=1}^m \frac{(2\pi)^{2n} B_n}{(2n)!} t^{2n-1} + O(t^{2m+1}) \right] dt. \quad (\text{A} \cdot 6)$$

Using (A.3) again, we have

$$= \frac{1}{2i \sin \pi s} \left[\frac{1}{\Gamma(s+1)} (e^{-\pi i} \log z)^s - \frac{1}{2\Gamma(s)} (e^{-\pi i} \log z)^{s-1} - \sum_{n=1}^m \frac{(2\pi)^{2n} B_n}{\Gamma(s+1-2n) (2n)!} (e^{-\pi i} \log z)^{s-2n} + O((\log z)^{s-2m-2}) \right]. \quad (\text{A} \cdot 7)$$

In the same way

$$I_1 = \frac{1}{2i \sin \pi s} \left[-\frac{1}{\Gamma(s+1)} (e^{\pi i} \log z)^s - \frac{1}{2\Gamma(s)} (e^{\pi i} \log z)^{s-1} + \sum_{n=1}^m \frac{(2\pi)^{2n} B_n}{\Gamma(s+1-2n) (2n)!} (e^{\pi i} \log z)^{s-2n} + O((\log z)^{s-2m-2}) \right]. \quad (\text{A} \cdot 8)$$

Hence

$$\phi(z, s) = -\frac{(\log z)^s}{\Gamma(s+1)} + \Gamma'(1-s) (\log z)^{s-1} + \sum_{n=1}^m \frac{(2\pi)^{2n} B_n}{\Gamma(s+1-2n) (2n)!} (\log z)^{s-2n} + O((\log z)^{s-2m-2}). \quad (\text{A} \cdot 9)$$

where $0 \leq \arg z < 2\pi$.

(A.9) without the second term coincides with the principal value of the Bose integral obtained by Clunie⁽¹⁵⁾ in case z is real and $z > 1$.

* We use the definition in Whittaker and Watson, *Modern Analysis*

From the definition of $\phi(z, s)$ and (A.9)

$$\begin{aligned}\phi(e^{\pi i} z, s) &= \phi(z^2, s) / 2^{s-1} - \phi(z, s) \\ &= -\frac{(\log z)^s}{\Gamma(s+1)} - \sum_{n=1}^m (1-2^{1-2n}) \frac{(2\pi)^{2n} B_n}{\Gamma(s+1-2n) (2n)!} (\log z)^{s-2n} \\ &\quad + O((\log z)^{s-2m-2}).\end{aligned}\tag{A.10}$$

When z is real, (A.10) agrees with Sommerfeld's¹³⁾ result derived by the partial integration. (A.10) is appropriate for evaluating $\phi(z, s)$ when real part of z is negative.

References

- 1) F. London, Phys. Rev. **54** (1938) 947. *Superfluids* II, New York, (1954).
- 2) R. H. Fowler and H. Jones, Proc. Cambr. Phil. Soc. **34** (1938), 573.
- 3) S. R. de Groot, J. G. Hooyman, and C. A. ten Seldam, Proc. Roy. Soc. A **203** (1950), 266.
- 4) P. T. Landsberg, Proc. Cambr. Phil. Soc. **50** (1954), 65.
- 5) A. R. Frazer, Phil. Mag. (7) **42** (1951), 156, 165.
- 6) G. Schubert, Z. Naturforsch. **1** (1946), 113. **2a** (1947), 250.
- 7) R. B. Dingle, Proc. Cambr. Phil. Soc. **45** (1949), 275.
- 8) S. Katsura and H. Fujita, Prog. Theor. Phys. **6** (1951), 498. Jour. Chem. Phys. **19** (1951), 795.
- 9) S. Katsura, Prog. Theor. Phys. **13** (1955), 571.
- 10) K. Ikeda, Proc. Int. Conf. Theor. Phys. p. 544. Kyôto (1953). Busseironkenkyû, No. 65 (1953), 145.
- 11) E. Lindelöf, *Le calcul des Résidus*, p. 138 Paris, (1905).
- 12) C. Truesdell, Ann. Math, **46** (1945), 144.
- 13) A. Sommerfeld, Zeits. f. Phys. **47** (1928), 1.
- 14) J. E. Robinson, Phys. Rev. **83** (1951), 678.
- 15) J. Clunie, Proc. Phys. Soc. A **67** (1954), 632.
- 16) C. N. Yang and T. D. Lee, Phys. Rev. **87** (1952), 404, T. D. Lee and C. N. Yang, Phys. Rev. **87** (1952), 410.
- 17) B. Widom, Phys. Rev. **96** (1954), 16.
- 18) H. S. Shapiro, Phys. Rev. **99** (1955), 1673.

Errata

Phase Transition of Husimi-Temperley Model of Imperfect Gas

Shigetoshi KATSURA

(Prog. Theor. Phys. **13** (1955), 571.)

| | | | | |
|--------|---------|--|-------------|---|
| p. 580 | line 7 | $z = \rho \exp(-\sum_{k=1}^{\infty} \beta_k^0 \rho^k)$ | should read | $z = \rho x^{-4} \exp(-\sum_{k=1}^{\infty} \beta_k^0 \rho^k)$ |
| p. " | line 10 | condensation point | " | temperature T_m |
| p. " | line 11 | temperature T_m | " | condensation point |

Two-Nucleon Problem with Pion Theoretical Potential, III — p - p Scattering at 18.2 Mev —

Junji IWADARE[†], Shoichiro OTSUKI,^{††} Ryozo TAMAGAKI^{††}

and

Wataro WATARI^{††}

Department of Physics, Kyoto University, Kyoto^{†,††}

Department of Physics, Hiroshima University, Hiroshima^{†††}

(Received July 7, 1956)

p - p scattering experiment at 18 Mev is analysed from the point of view of the pion theory. The pion theoretical potential, which is consistent with all low energy phenomena, gives a precise fit to the experimental data at 18 Mev. Especially, it is found that in the triplet odd state an attractive central potential around the pion range is required from the experimental data. This attractive potential is consistent with the two-pion-exchange potential.

§ 1. Introduction

In the present work, p - p scattering experiment at 18.2 Mev is analysed from the point of view of the pion theory.

From this point of view, two-nucleon problems at low energies were discussed by the present authors: The deuteron problem in the paper I¹, the singlet even state in II²⁾ and the triplet odd state³ in the paper entitled as "Meson Theoretical Potentials in Triplet Odd State". In the series of these works, the one-pion-exchange potential, which is the first and the most natural consequence of the pion theory, was quantitatively established. The effective coupling constant of the one-pion-exchange potential was determined as $g_\pi^2/4\pi = 0.080 \pm 0.010$ and all low energy phenomena were explained consistently by the pion theory of nuclear forces.

One might expect that he could gain more and more essential knowledge of nuclear forces by analysing the intermediate and high energy phenomena. However, from the pion theoretical point of view, they are not so suitable for a detailed investigation into nuclear forces as it is usually believed. Many parameters — phase shifts and mixing ratios of the two components of the eigenwaves of scattering — are included in the expression of the scattering cross section. About a half of the parameters are expected to be affected more or less by the quantitatively unreliable potential in the region II ($0.7 \lesssim x \lesssim 1.5$) and by the little known interactions in the region III ($x \lesssim 0.7$). Therefore, what we can expect to obtain by analysing the intermediate and high energy phenomena is, at

[†] Now at the Research Institute for Fundamental Physics, Kyoto University, Kyoto.

the present stage of the pion theory, not quantitative information but only qualitative one. Since we know the quantitative properties of the one-pion-exchange potential in the outer region, some qualitative information on the two-(or more-)pion-exchange potential can be obtained by making full use of the known properties of the one-pion-exchange potential.

As is seen from the above discussion, it seems useful to analyse an accurate experiment of nucleon-nucleon scattering at not so high energy. For this purpose we have chosen the p - p scattering experiment at 18.2 ± 0.2 Mev made at Princeton¹⁾. The new information we can expect to obtain from it is mainly concerned with the two-pion-exchange potential which is dominant in the region II ($0.7 \lesssim x \lesssim 1.5$), since the impact parameter* of the P -wave, which is mainly responsible to the angular distribution, is nearly equal to 2 at 18 Mev. The same experiment has already been analysed by some of the present authors and their collaborator, i.e. by Otsuki and Fujii⁶⁾ and by Otsuki and Tamagaki⁷⁾, and by other people, i.e. by Martin and Verlet⁸⁾ and by Beretta, Clementel and Villi⁹⁾. However, it is to be emphasized that the present analysis is based on the validity of the one-pion-exchange potential in the outer region in contrast to other phenomenological analyses. Other experiments on nucleon-nucleon scattering at a few tens Mev, especially the n - p scattering, are not so accurate. The photodisintegration of the deuteron at the photon energy of about 20 Mev may serve to elucidate some aspects of nuclear forces, and will be discussed in the forthcoming paper IV.¹⁰⁾

We describe the method and the results of our analysis in § 2 and § 3 respectively. The discussion and conclusion are given in § 4.

The results support the pion theory of nuclear forces, since the theory which is consistent with all low energy phenomena can also give a precise fit to the data at 18.2 Mev. In particular, the data require an attractive central force in the region II in the triplet odd state and this attractive force is consistent with the prediction of the two-pion-exchange potential and with the result of the low energy p - p scattering.

§ 2. Method

Our procedure for analysing the 18.2 Mev p - p scattering data is as follows:

To assume the 1S_0 -wave phase shift ${}^1\delta_0$ as a free parameter.

To determine the 1D_2 -wave phase shift ${}^1\delta_2$ by the one-pion-exchange potential.

To calculate the ${}^3P_{0,1,2}$ -wave phase shifts ${}^3\delta_0^1$, ${}^3\delta_1^1$, ${}^3\delta_2^1$ (the suffices 0, 1, 2 denoting the total angular momentum) using the following potential

$$\begin{aligned}
 V &= (1/3) (g_e^2/4\pi) \mu c^2 (1 + S_{12} (1 + 3/\kappa + 3/\kappa^2)) e^{-\kappa}/\kappa \text{ for } \kappa > 1, \\
 &\quad \text{(the one-pion-exchange potential)} \quad (1) \\
 &= V_{00} + S_{12} V_{0T} \text{ (} V_{00} \text{ and } V_{0T} \text{ are two constant.) for } \kappa < 1.
 \end{aligned}$$

*) The impact parameter b is defined as $b\kappa = \sqrt{l(l+1)}$.⁵⁾

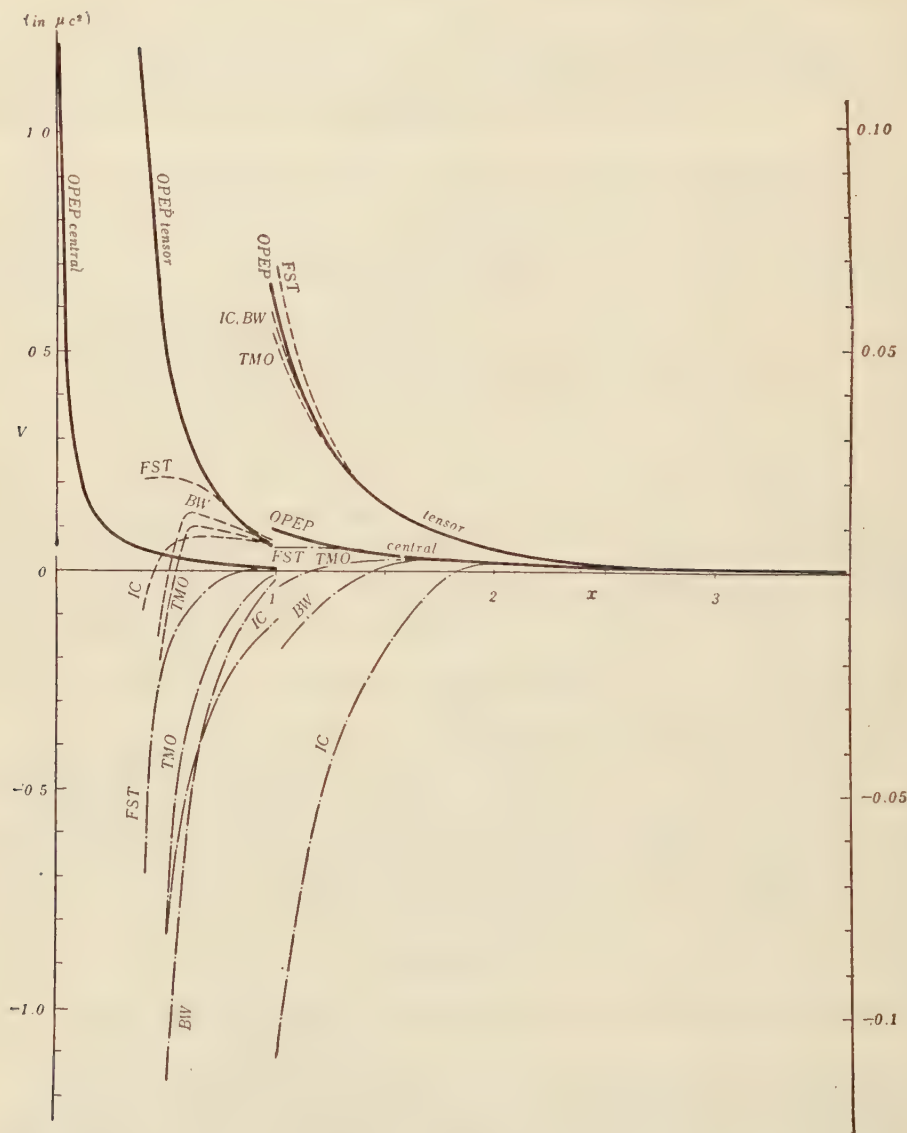


Fig. 1. The pion theoretical potential in the triplet odd state.

OPEP: The one-pion-exchange potential with $g_\pi^2/4\pi=0.08$.

FST: The potential constructed by Fukuda, Sawada and Taketani's method, with $g_\pi^2/4\pi=0.08^{(1)(a)(11b)}$. The probability that the nucleons are to be bare is properly taken into account. The high frequency part of the pion field is cut off by the Gaussian factor with the cut-off momentum $\hbar k_c=6\mu c$.

TMO: The g^2+g^4 potential in the perturbation expansion calculated by Taketani, Machida and Onuma, with $g^2/4\pi=0.08^{(1)(c)}$. The probability part is also expanded in powers of the coupling constant.

BW: The g^2+g^4 potential calculated by Brueckner and Watson with $g^2/4\pi=0.08^{(1)(d)}$. The probability part is approximated by unity.

IC: The one- and two-pion-exchange potential in the intermediate coupling theory with $g_\pi^2/4\pi=0.08^{(1)(e)}$. The cut-off momentum of the Gaussian cut-off factor is $\hbar k_c=4.1\mu c$.

The coupling between the 3P_2 - and 3F_2 -waves are neglected. The reason for the above procedure is as follows:

We have examined the singlet even state potential qualitatively in the work II²⁾ and found that it is consistent with the prediction of the pion theory. However, if we calculate ${}^1\delta_0$ using this potential the value of this phase shift at 18 Mev is uncertain in a fairly wide range, because the inner potential has a large ambiguity. Therefore we should treat ${}^1\delta_0$ as a free parameter and adjust it to reproduce the cross section at $\theta=90^\circ$.

On the contrary, ${}^1\delta_2$ is almost completely determined by the one-pion-exchange potential, since the impact parameter of the D -wave is as large as 3.7 at 18 Mev.

The P -wave phase shifts at 18 Mev are expected to be mainly determined by the one-pion-exchange potential, since the impact parameter is 2.2. However, in contrast with the D -wave phase shift, they are affected also by the two-pion-exchange potential. Since we do not know the exact shape of the two-pion-exchange potential as can be seen from Fig. 1, we take into account the effect of the two-pion-exchange potential by introducing a phenomenological square well potential in the inner region, i.e., $V_{OC}+S_{12}V_{OT}$ of Eq. (1).

From the present view point of the theory of nuclear forces some results of the previous works by two of the authors are out of date:^{(6),(7)} The most serious fault in these works is that the 1S_0 -wave phase shift was calculated with the g^2+g^4 potential in the perturbation approximation^(11c) everywhere in the region $\kappa>0.6$, and then its effects was subtracted from the experimental data in order to identify the remainder as the triplet scattering. At that time, the g^2+g^4 potential in the singlet even state was regarded much more reliable than that in the triplet odd state. Now the situation has been much changed with respect to the reliability of the pion theoretical potential owing to the new development, as was described in § 1 of our work I.¹⁾

§ 3. Results

The value of the singlet D -wave phase shift is calculated to be ${}^1\delta_2=0.25^\circ$ for $g_e^2/4\pi=0.08$. Since the potential at small distances has very little influence on the phase shift, the value is expected in the first approximation to be proportional to $g_e^2/4\pi$, the strength of the one-pion-exchange potential.

Next, we assume three values for V_{OT} , the depth of the inner tensor potential in the triplet odd state, as

$$\begin{aligned} V_{OT} &= 0 && \text{(zero cut-off),} \\ &= (7/3) (g_e^2/4\pi) \mu c^2 e^{-1} && \text{(straight cut-off of the one-pion-exchange} \\ &&& \text{potential at } \kappa=1), \\ &= 2 \times (7/3) (g_e^2/4\pi) \mu c^2 e^{-1} && \text{(twice the straight cut-off).} \end{aligned}$$

Such a choice of V_{OT} is sufficient for the analysis of the 18 Mev data from our point of view since the two-pion-exchange potential is expected to be small (about 10% even at $\kappa=1$) and consequently the one-pion-exchange potential is the main part of the tensor force around and outside the pion range (see Fig. 1).

The situation is complicated for the central force in the triplet odd state, as the potential due to the two-pion-exchange processes is expected to surpass the one-pion-exchange potential outside the pion range (see Fig. 1). Indeed the g^4 potential in the perturbation approximation is a few times stronger than the g^2 potential at $x=1$. Furthermore their signs are different. Therefore, we change the depth of the inner central potential V_{OT} over a wide range, i.e., from 0 to -50 Mev.

The value of the adjustable parameter ${}^1\delta_0$, the singlet S-wave phase shift, is determined so as to reproduce $d\sigma(90^\circ)/d\Omega$ in each case.

The result for the case of $g_e^2/4\pi=0.08$ is tabulated in Table I. We can find from the table that when V_{OT} is large, ${}^1\delta_0$ is comparatively small. The reason is that the strong (positive) tensor potential gives a large isotropic cross section of the triplet scattering, which excludes the large cross section due to the singlet scattering.

On the contrary, the angular distribution around $\theta\sim 30^\circ$, where the interference effect between the nuclear and the Coulomb scatterings is appreciable, is sensitive to V_{OC} as is expected. Consequently the upper and lower limits for V_{OC} can be derived from the angular distribution, in particular from the difference $d\sigma(90^\circ)/d\Omega - d\sigma(30^\circ)/d\Omega$.

Table I. Phase shifts and the p - p angular distribution at 18.2 Mev calculated according to the procedure of § 2. The coupling constant of the one-pion-exchange potential is taken as $g_e^2/4\pi=0.08$. Experimental data are

$$d\sigma(90^\circ)/d\Omega - d\sigma(30^\circ)/d\Omega = 2.32 \pm 0.4 \text{ mb,}$$

$$\text{and } C_2 < 0.01$$

where C_2 is the coefficient of $\cos^2\theta$ in the triplet nuclear scattering. $d\sigma(90^\circ)/d\Omega$ can be reproduced since ${}^1\delta_0$ is adjusted. Some examples of the angular distribution given by the phase shifts of this table are shown in Fig. 2.

| V_{OT} | zero cut-off | | straight cut-off | | twice the straight cut-off | |
|---|--------------|-------|------------------|-------|----------------------------|-------|
| V_{OC} (Mev) | 0 | -50 | 0 | -50 | 0 | -50 |
| ${}^3\delta_0^1$ (deg.) | 8.4 | 11.2 | 9.8 | 16.7 | 13.7 | 27.9 |
| ${}^3\delta_1^1$ (deg.) | -3.7 | -2.9 | -3.9 | -3.3 | -4.1 | -3.6 |
| ${}^3\delta_2^1$ (deg.) | 0.1 | 1.3 | 0.2 | 1.4 | 0.3 | 1.6 |
| ${}^1\delta_0$ (deg.) | 51.1 | 50.6 | 50.6 | 48.3 | 49.6 | 40.6 |
| ${}^1\delta_2$ (deg.) | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| $d\sigma(90^\circ)/d\Omega$ $-d\sigma(30^\circ)/d\Omega$ (mb.) | 0.63 | 2.2 | 0.8 | 2.7 | 1.4 | 3.2 |
| C_2 | 0.003 | 0.007 | 0.004 | 0.009 | 0.005 | 0.018 |

By linear interpolation from Table I we can estimate the values of V_{OC} and ${}^1\delta_0$ which fit to the experimental data. For the three assumed values of V_{OT} they are as follows:

| V_{OT} | V_{OC} (in Mev) | ${}^1\delta_0$ (in degree) |
|----------------------------|-------------------|----------------------------|
| zero cut-off | $-40 \sim -65$ | $50.7 \sim 50.4$, |
| straight cut-off | $-28 \sim -50$ | $50.4 \sim 48.3$, (2) |
| twice the straight cut-off | $-14 \sim -37$ | $47.0 \sim 43.3$. |

Therefore in the case of $g_e^2/4\pi=0.08$, it may be reasonable to conclude that

$$V_{OC} = -10 \sim -70 \text{ Mev and } {}^1\delta_0 = 51^\circ \sim 43^\circ.$$

When V_{OC} is vanishingly small or positive, the effect of the repulsive central part of the one-pion-exchange potential becomes dominating. In such a case the interference minimum around $\theta \sim 30^\circ$ becomes less pronounced as the curve III in Fig. 2. On the contrary, when V_{OC} is about -50 Mev, this inner attractive central potential diminishes the effect of the repulsive central part of the one-pion-exchange potential. In such a case, the interference minimum becomes quite appreciable. This can be understood when one notices that the minimum is caused by the destructive interference between the Coulomb scattering and the P -wave nuclear scattering, where the effects of the tensor force to the latter are likely to cancel out one another.

Calculation was also made for other values of $g_e^2/4\pi$ ($=0.07$ and 0.09). The adjusted values of V_{OC} are scarcely changed, i.e., $V_{OC}=0 \sim -70$ Mev. For $g_e^2/4\pi=0.07$ the value of the adjusted 1S_0 -wave phase shift ${}^1\delta_0$ increases by about $1 \sim 2^\circ$ from the corresponding value for $g_e^2/4\pi=0.08$. For $g_e^2/4\pi=0.09$, it decreases by almost the same amount.

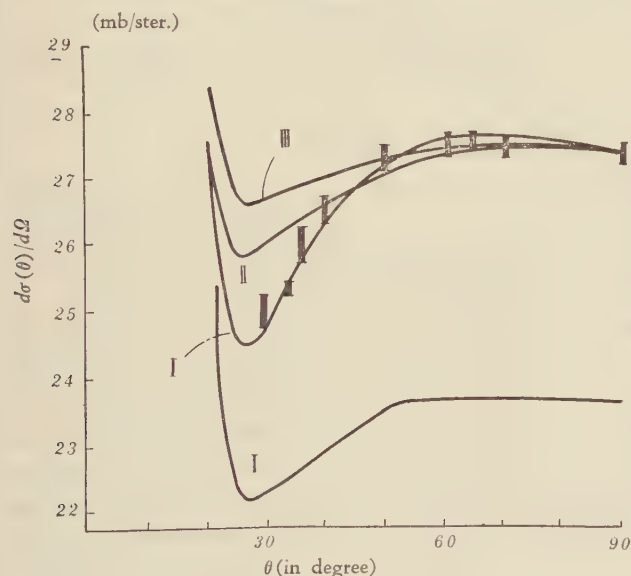


Fig. 2. Theoretical curves of the p - p scattering at 18.2 Mev.

| | ${}^1\delta_0$ | ${}^1\delta_2$ | ${}^3\delta_0^1$ | ${}^3\delta_1^1$ | ${}^3\delta_2^1$ |
|-----|----------------|----------------|------------------|------------------|------------------|
| I | 48.3 | 0.25 | 16.7 | -3.3 | 1.4 |
| I' | 48.3 | 0.25 | — | — | — |
| II | 54.2 | 0.25 | — | — | — |
| III | 50.7 | 0.25 | 9.8 | -3.9 | 0.2 |

Curves I and III are due to the pion theoretical potential of Eq. (1) with $g_e^2/4\pi=0.08$. V_{OT} is assumed as the straight cut-off of the one-pion-exchange potential at $x=1$. $V_{OC}=-50$ Mev for I and $V_{OC}=0$ for III respectively.

Curves I' and II are the angular distributions when no interaction is present in the triplet odd state.

Experimental data are those of reference 4).

Note: The lowest curve I in the figure should be read as I'.

§ 4. Discussion and conclusion

In the previous section we have obtained the result that the inner central potential in the triplet odd state of Eq. (1) is

$$V_{OC}=0 \sim -70 \text{ Mev.} \quad (3)$$

Now it is very interesting to point out that the value of this inner central potential is in agreement with that estimated from the negative triplet P -wave phase shift at low energies.^{3),7)}

The repulsive one-pion-exchange central potential in the triplet odd state gives a negative P -wave phase shift at low energies, which is, however, somewhat larger than the experimental one in its magnitude, if $g_c^2/4\pi$ of the one-pion-exchange potential is larger than about 0.06. Therefore an attractive interaction must be present in the inner region. If this attractive interaction is represented by the potential of the shape of Eq. (1), the depth V_{oc} estimated from the negative P -wave phase shift at low energies is $V_{oc}=0\sim-50$ Mev. Thus the results of the two independent estimates which were made in two different energy regions agree with each other surprisingly well. It is very likely that this agreement is due to our correct choice of the coupling constant $g_c^2/4\pi\sim 0.08$ for the one-pion-exchange potential.

Of course, the attractive potential V_{oc} is consistent with the prediction of the two-pion-exchange potential, not only in its tendency but also in its magnitude as can be seen from Fig. 1.

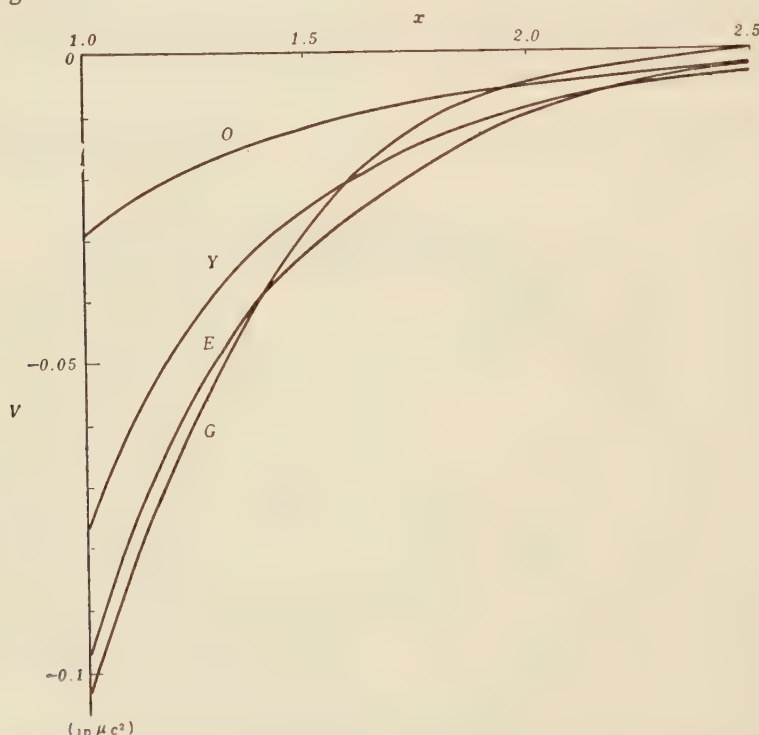


Fig. 3. The outer part of potentials in the singlet even state. The potentials can reproduce the low energy experimental data that $^1a=-20\times 10^{-13}$ cm and $^1r_e=2.5\times 10^{-13}$ cm.

O: the one-pion-exchange potential with $g_c^2/4\pi=0.08$,

Y: the Yukawa potential,

E: the exponential potential,

G: the Gauss potential.

Notice that the one-pion-exchange potential is weaker in the outer region than the long-tailed phenomenological potentials Y and E.

Here it is worth-while to emphasize the conclusion previously obtained with respect to the D -wave phase shift ${}^1\delta_2$.⁽⁶⁾ ${}^1\delta_2$ is almost completely determined by the one-pion-exchange potential and is rather small, since the one-pion-exchange potential with $g_c^2/4\pi = 0.080 \pm 0.010$ is weaker than usual phenomenological potentials in the outer region (see Fig. 3). The small ${}^1\delta_2$ is very suitable to reproduce the isotropic angular distribution, since otherwise the interference between the 1S_0 - and 1D_2 -waves gives a deep minimum at 90° leading to a serious discrepancy.

Summarizing the above discussions, we can conclude that the p - p scattering experiment at 18.2 Mev supports the pion theoretical potential in the following points:

- (1) The one-pion-exchange potential gives the small 1D_2 -wave phase shift.
- (2) An attractive central potential is to be added in the inner region to the one-pion-exchange potential in the triplet odd state. This attractive potential is consistent with the prediction of the two-pion-exchange potential.
- (3) A striking agreement is obtained for the depths V_{00} of the above inner central potential estimated by two analysis of the intermediate and low energy p - p scattering experiments.

An example of the phase shifts due to the pion theoretical potential is given below. It gives a precise fit to the experimental data:

$$\begin{aligned} {}^1\delta_0 &= 48.3^\circ, & {}^1\delta_2 &= 0.25^\circ, \\ {}^3\delta_0^1 &= 16.7^\circ, & {}^3\delta_1^1 &= -3.3^\circ, & {}^3\delta_2^1 &= 1.4^\circ, \end{aligned} \quad (4)$$

where $g_c^2/4\pi$ of the one-pion-exchange potential is 0.08; the depths of the inner potential of Eq. (1) are assumed as $V_{00} = -50$ Mev and $V_{0T} = 0.08(7/3e)\mu c^2$ (the straight cut-off at $x=1$); ${}^1\delta_0$ is adjusted to reproduce the cross section at 90° .

Though we could determine the upper and lower bounds of the singlet S -wave phase shift ${}^1\delta_0$, we would not enter into its detailed investigation. It will be clear that we can not get meaningful results from such investigation at the present stage, since the singlet S -wave shifted by ${}^1\delta_0$ in Table I has its maximum just around the pion range and is strongly affected by the inner potential. If the exact shape of the two-pion-exchange potential were established, we could obtain from ${}^1\delta_0$ some knowledge of the interaction at small distances, for example, by solving the Schrödinger equation with the one- plus two-pion-exchange potential starting from outside with the boundary condition that the wave is shifted by ${}^1\delta_0$.

Martin and Verlet⁽⁸⁾ also calculated the angular distribution using Lévy's potential⁽¹²⁾. Their good fit is not surprising, as Lévy's potential with $G^2/4\pi = 10.36$ is nothing but the one-pion-exchange potential with $g_c^2/4\pi = 0.057$ plus an attractive central potential of a short range ($\sim \exp(-2x)$).

Recently, there was an attempt to determine the phase shifts phenomenologically from the experimental data by Beretta, Clementel and Villi⁽⁹⁾. It seems to us, however, that there might be too large ambiguity in the result if one takes into account the experimental error. At such an energy, we think, an analysis on the standpoint of the pion theory may be much more superior than the purely phenomenological phase shift

analysis.

However, the following fact may be worth noticing: A simplified phase shift analysis for the present experimental data assuming only a central force gives the result that ${}^1\delta_0 = 54.1^\circ$, ${}^3\delta^1$ (3P -wave phase shift due to a central force) $= +1.0^\circ$ and ${}^1\delta_2 = +0.4^\circ$.⁴⁾ At low energies ($E < 5$ Mev), ${}^3\delta^1$ is negative, which is the consequence of the repulsive character of the central one-pion-exchange potential.³⁾ Thus it is found that ${}^3\delta^1$ changes its sign as the energy goes higher. This is just what is expected from the pion theory. As the energy goes higher, the attractive central force of the two-pion-exchange potential begins to influence ${}^3\delta^1$. The strong tensor potential also tends to make ${}^3\delta^1$ positive as if an attractive potential were added to the central potential. Therefore when a simplified phase shift analysis neglecting the tensor force is applied to the experimental data of p - p scattering, the resultant ${}^3\delta^1$ is expected from the pion theory to change its sign around 10 Mev.

The authors wish to thank Prof. M. Taketani and Prof. T. Toyoda for their discussions and encouragement.

References

- 1) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Prog. Theor. Phys.* **16** (1956), 455.
- 2) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Prog. Theor. Phys.* **16** (1956), 472.
- 3) S. Otsuki and R. Tamagaki, *Prog. Theor. Phys.* **12** (1954), 806.
- 4) J. L. Yntema and M. G. White, *Phys. Rev.* **95** (1954), 1226.
- 5) See, for example, J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Supplement of the Progress of Theoretical Physics*, No. 3 (1956); Part II, A-5.
M. Matsumoto and W. Watari, *Prog. Theor. Phys.* **11** (1954), 63.
- 6) S. Otsuki and S. Fujii, *Prog. Theor. Phys.* **12** (1954), 521.
- 7) S. Otsuki and R. Tamagaki, *Prog. Theor. Phys.* **14** (1955), 52.
- 8) A. Martin and L. Verlet, *Phys. Rev.* **89** (1954), 519.
- 9) L. Beretta, E. Clementel and C. Villi, *Nuovo Cimento* **1** (1955), 739.
- 10) J. Iwadare, S. Otsuki, M. Sano, S. Takagi and W. Watari, *Prog. Theor. Phys.* **16** (1956), 641.
- 11a) N. Fukuda, K. Sawada and M. Taketani, *Prog. Theor. Phys.* **12** (1954), 156.
- 11b) K. Inoue, S. Machida, M. Taketani and T. Toyoda, *Prog. Theor. Phys.* **15** (1956), 122.
- 11c) M. Taketani, S. Machida and S. Onuma, *Prog. Theor. Phys.* **7** (1952), 45.
- 11d) K. A. Brueckner and K. M. Watson, *Phys. Rev.* **92** (1953), 1023.
- 11e) Y. Nogami, and H. Hasegawa *Prog. Theor. Phys.* **15** (1956), 137.
- 12) M. M. Lévy, *Phys. Rev.* **87** (1952), 725.
- 11) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, *Supplement of the Progress of Theoretical Physics*, No. 3 (1956).

Formation of Light Nuclei in the Expanding Universe

Chushiro HAYASHI and Minoru NISHIDA

Department of Physics, Kyoto University, Kyoto

(Received September 4, 1956)

The building up of light nuclei up to Ne^{20} out of initial protons and neutrons in the early stage of the expanding universe is studied, taking account of the α -capture reactions. If we take the matter density $\rho_m = 10^6 \sim 10^7 \text{ g/cm}^3$ at $T = 10^{10} \text{ }^\circ\text{K}$, we can obtain the relative abundance of hydrogen, helium and oxygen, consistent with present observation. The possibility of forming heavier nuclei is discussed.

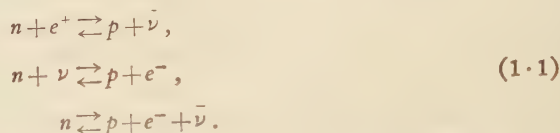
§ 1. Introduction and summary

One of the outstanding astrophysical problems is the origin of chemical elements in our universe. Two different hypotheses have been proposed so far. In one the elements are assumed to be formed during the pre-stellar stage of the universe. The so-called α - β - γ theory¹⁾ is a typical one based on this hypothesis. According to this theory the primordial substance from which the elements were formed is assumed to have been a highly compressed neutron gas. As the gas temperature fell as the result of the expansion of the universe, neutrons decayed into protons, and the elements were built up through successive neutron captures and β -decays. The most serious difficulty of the α - β - γ theory is the fact that a sufficient amount of carbon and oxygen could not have been formed due to the absence of stable nuclei of mass number 5 and 8.²⁾

The other hypothesis, as suggested by Hoyle,³⁾ Fowler and Greenstein,⁴⁾ is that the elements originated in dense and hot interior of stars, starting from protons only. This theory is supported by the anomalous abundances of the heavy nuclei observed on some stars. It also has the difficulties that the natural α -radioactive nuclei will not be formed through nuclear reactions which proceed in a rather long time scale under the stellar condition, and that we have no observational evidence for the stars composed entirely of hydrogen. It must be admitted, however, that some fraction of the present matter has undergone nuclear reactions in its stellar stage.

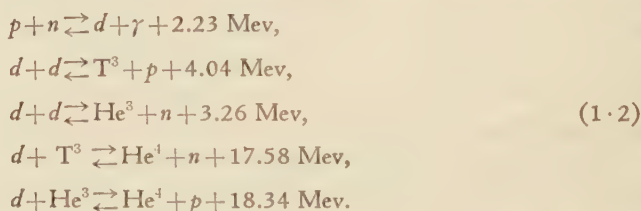
At present, we can not assert which idea, pre-stellar or stellar, is preferable for explaining the main feature of chemical abundance. This problem will be intimately connected with the further development of the evolutionary scheme of stars as emphasized by Taketani, Hatanaka and Obi.⁵⁾ In this paper we shall examine whether or not the above mentioned difficulties of the α - β - γ theory can be avoided by considering the α -capture processes, such as $3\text{He}^1 \rightarrow \text{C}^{12}$, which have recently been studied by Hayakawa et al.⁶⁾ using the new experimental data.

For the formation of C^{12} nuclei by the α -reactions, it is necessary to take the matter density to be much higher than in the case of the α - β - γ theory. Under such a condition the expansion of the universe will be faster than the neutron-decay. On this point it is favorable to follow the improvement made by one of the present authors^{7)*} that there exist protons several times more abundant than neutrons in the pre-elements stage as the results of their interactions with energetic electrons and neutrinos such as



The neutron-proton ratio at the beginning of the formation of light nuclei was found to be nearly 1:6 in the case of low matter densities as used in the α - β - γ theory. Applying the results in I to the case of high matter densities $\rho_m = 10^7 \sim 10^8 \text{ g/cm}^3$ at the temperature $10^{10} \text{ }^\circ\text{K}$, we find that the n - p ratio is 1:2.5~1:4 as shown in § 2.

Now we consider the formation of light nuclei, dividing it into two stages. In the first stage, very light nuclei up to He^4 are formed out of the above neutrons and protons and in the second stage carbon, oxygen and heavier nuclei are built up. The first stage begins at $T \lesssim 9 \times 10^9 \text{ }^\circ\text{K}$ in the case under consideration. When the expansion of the universe proceeds and its temperature falls to the one corresponding to the nuclear binding-energies, neutrons and protons react with each other to form deuterons, tritons, He^3 and He^4 . Among these reactions we choose the following processes with high reaction rates:



In the state of such high temperature, $T \sim 9 \times 10^9 \text{ }^\circ\text{K}$, we must include the inverse reactions but can neglect the β^2 decays because of the fast expansion of the universe. Since all the nuclear reactions, (1.2), occur faster than the expansion of the universe at the beginning of this stage, statistical equilibrium will exist. As the temperature falls, the rates of inverse reactions with large Q values decrease, which produce a deviation from the equilibrium state. Calculations show that, at the end of the first stage, we have

$$n_p : n_{\text{He}^4} = 3 : 1 \sim 6 : 1,$$

and the numbers of neutrons, deuterons, tritons and He^3 nuclei are very small ($< 10^{-7}$

* Hereafter referred to as I. In I, the calculation was performed using a wrong value of neutron decay life-time, 20 min. instead of 12.8 min. due to the uncertain experimental data at that time. When the correct value is used, the numerical results in I agree with those by Alpher, Follin and Herman.⁸⁾

that of protons), as shown in Figs. 2a and 2b in § 3. Such a small amount of neutrons will be insufficient to build up the heavy elements according to the scheme of the α - β - γ theory. As for the later stage, we have only to consider the reactions in which protons and He^4 play a role.

When the temperature falls to $4 \sim 5 \times 10^9$ °K, the formation of C^{12} and O^{16} by successive α -capture reactions begins as the second stage. In our case the matter density is so high that a considerable amount of Be^8 nuclei can be in statistical equilibrium with He^4 .⁽⁶⁾ These Be^8 nuclei collide with other He^4 to form C^{12} nuclei which, in turn, capture additional He^4 to form O^{16} nuclei. Taking the matter density $\rho_m = 10^6 \sim 10^7$ g/cm³ at $T = 10^{10}$ °K, we obtain a final O^{16} abundance in accordance with the observational cosmic abundance, but the calculated C^{12} - O^{16} ratio is comparatively smaller than the observational one. As a possibility to avoid this difficulty we can consider the spallation of O^{16} nuclei by non-thermal protons which are accelerated by the turbulence accompanying the expansion of the universe (see in § 4).

Finally, in § 5, we will discuss the results of our calculations as well as the other possible routes of forming C^{12} and O^{16} out of only protons and He^4 nuclei. As the next stage of the formation of light nuclei, we shall have to consider the processes forming Ne^{20} and heavier nuclei from protons, He^4 and O^{16} nuclei. In this stage, the expansion of the universe is so slow that β -decays, such as $\text{Na}^{21} \rightarrow \text{Ne}^{21} + e^+$, $\text{Al}^{25} \rightarrow \text{Mg}^{25} + e^+$, etc., can not be neglected. As a result of (α, n) reactions involving Ne^{21} , Mg^{25} , etc., some neutrons will be produced. Therefore it may be possible that heavy elements are synthesized by the captures of these neutrons.

§ 2. Proton-neutron ratio before the formation of He^4 nuclei

According to the general theory of relativity, the expansion of the universe in the stage of high density is given by

$$\frac{1}{l} \frac{dl}{dt} = \left(\frac{8\pi}{3} G \rho \right)^{1/2}, \quad (2.1)$$

where l is an arbitrary proper length of a volume containing a given amount of matter, ρ the total mass density and G the gravitational constant. We restrict ourselves to the case in which the matter density, ρ_m , is greater than the radiational mass density $\rho_r = aT^4/c^2$, that is,

$$\rho_m > 10^5 \text{ g/cm}^3 \quad \text{at } T = 10^{10} \text{ °K}.$$

If $\rho_m < 10^5 \text{ g/cm}^3$ at $T = 10^{10}$ °K, radiation pressure P_r is greater than matter pressure P_m . Further, in the case in which $\rho_m \lesssim 10^7 \text{ g/cm}^3$ at $T = 10^{10}$ °K, we can neglect the rather complicated effect of the degeneracy of electrons on the reactions (1.1) as shown in I.* For simplicity we take as the range of ρ_m

* According to (9) of I, we find λ , which expresses the degree of degeneracy, to be less than 0.2 in the case under consideration.

$$10^5 \lesssim \rho_m \lesssim 10^7 \text{ g/cm}^3 \quad \text{at } T=10^{10} \text{ }^\circ\text{K.} \tag{2.2}$$

It will be shown in § 4 that an appropriate amount of oxygen can be formed under the condition (2.2). Neglecting the contribution of kinetic energy of matter to ρ , we obtain from the equation of adiabatic expansion,

$$d(\rho l^3)/dl^3 = d(\rho_r l^3)/dl^3 = -\frac{1}{3} \rho_r. \tag{2.3}$$

From (2.3) we have $\rho_r \propto l^{-4}$ and $T \propto l^{-1}$. As $\rho_m \propto l^{-3}$ we can put

$$\rho_m = \rho_0 (T/10^{10})^3, \tag{2.4}$$

with ρ_0 as a constant specifying the universe. We have from (2.1), (2.3) and (2.4)

$$\rho_m = 7.9 \times 10^5 t^{-2} \tag{2.5}$$

and

$$1/T = 1.08 \times 10^{-10} (\rho_0/10^6)^{1/3} t^{2/3}. \tag{2.6}$$

In Table 1 we show the time scale given by (2.6).

Table 1. The relation between t and T

| $T(10^9 \text{ }^\circ\text{K})$ | | 10 | 4 | 2 | 1 | 0.5 | 0.2 | 0.1 |
|----------------------------------|------------------------------|------|-----|-----|-----|-----|-----|-----|
| $t(\text{sec})$ | $\rho_0=10^6 \text{ g/cm}^3$ | 0.89 | 3.5 | 10 | 28 | 80 | 316 | 890 |
| | $\rho_0=10^7 \text{ g/cm}^3$ | 0.28 | 1.1 | 3.2 | 8.9 | 25 | 100 | 280 |

Under the above condition, (2.2), the proton-neutron ratio in the pre-elements stage can be calculated with slight modifications of I, that is, by using the correct neutron decay life-time 12.8 min. and by taking $\rho=\rho_m$ in (2.1). The calculated proton-neutron ratio is shown in Fig. 1.

Finally, we want to remark that in Newtonian mechanics we can use the same equation as (2.1) in the case when gas with spherically symmetrical distribution expands uniformly and when the magnitudes of its kinetic and potential energies are comparable. It is suggested that the supernovae explosions in their early stage are described approximately with the same time scale.

§ 3. Formation of He⁴ nuclei

Suppose a nuclear reaction among four kinds of nuclei A, B, C and D as

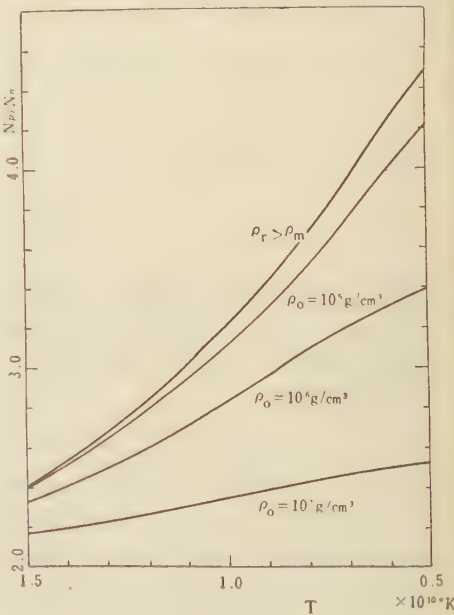
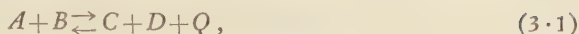


Fig. 1. Neutron-proton ratios for different values of ρ_0 . The case $\rho_r > \rho_m$ is added for comparison.



where Q stands for the energy release of the reaction. In our case, the nuclear reactions occur at so high temperature that in general their inverse reactions must be taken into account. Taking account of the expansion of the universe, the change in the number of nuclei C per unit volume, n_C , is expressed by

$$d(n_C/\rho)/dt = (1/\rho)[pn_A n_B - p'n_C n_D] \quad (3.2)$$

with

$$p/p' = n_A^{(0)} n_B^{(0)} / n_C^{(0)} n_D^{(0)}. \quad (3.3)$$

Here $n_A^{(0)}$ is the number of nuclei A in thermodynamical equilibrium at temperature T and density ρ :

$$n_A^{(0)} = g_A (2\pi H A k T / h^2)^{3/2} \exp[\{(A-Z)\mu + Z\lambda - M_A c^2\} / kT], \quad (3.4)$$

where A and Z are the mass number and atomic number of nucleus A , respectively; H is proton mass; λ , μ are the chemical potentials of proton and neutron; $g_A = 1 + 2I_A$ is the statistical weight of the ground state of nucleus A with spin I_A . The thermonuclear rate p in (3.2) is a function of temperature and is expressed as

$$p(T) = (8/\pi\mu)^{1/2} (kT)^{-3/2} \int_0^\infty \sigma(E) e^{-E/kT} E dE, \quad (3.5)$$

where $\mu = (AB/(A+B))H$ and $\sigma(E)$ is the reaction cross section, E being the kinetic energy in the center of mass system. It is convenient in what follows to use

$$w_i = n_i/n, \quad n = \rho/H \quad (3.6)$$

instead of n_i . Then, the time variation of w_C is given by

$$\frac{dw_C}{dt} = p(T) n \left\{ w_A w_B - \frac{g_A g_B}{g_C g_D} \left(\frac{AB}{CD} \right)^{3/2} e^{-Q/kT} w_C w_D \right\}. \quad (3.7a)$$

For the radiative capture, $A+B \rightleftharpoons C+\gamma$, we have in a similar way

$$\frac{dw_C}{dt} = p(T) n \left\{ w_A w_B - \frac{g_A g_B}{g_C} \left(\frac{AB}{C} \right)^{3/2} \left(\frac{2\pi H k T}{h^2} \right)^{3/2} e^{-Q/kT} w_C / n \right\}. \quad (3.7b)$$

If the expansion of the universe is slower than all the nuclear reactions in (1.2), the concentrations of all the nuclear species vary in accordance with equilibrium values corresponding to the temperature and density at each stage of the expansion. We can obtain equilibrium values $w_i^{(0)}$'s, as functions of temperature, from (3.4) and the conservation of the members of protons and neutrons

$$\begin{aligned} w_p + w_d + w_t + 2w_{\text{He}^3} + 2w_\alpha &= a, \\ w_n + w_d + 2w_t + w_{\text{He}^3} + 2w_\alpha &= b, \end{aligned} \quad (3.8)$$

where $a:b$ ($a+b=1$) is the proton-neutron ratio at the beginning of He^4 formation as calculated in § 2. The dotted curves in Figs. 2a and 2b show the variation of $w_i^{(0)}$'s obtained in this way.

As shown in Figs. 2a and 2b, the formation of He^4 begins at $T \lesssim 9 \times 10^9$ °K. In this stage the β decays can be neglected owing to the fast expansion of the universe, as seen from Table 1. We consider only the five reactions shown in (1.2), neglecting the others with much lower reaction rates at the temperatures under consideration. Reaction rates of (1.2) which are calculated from (3.5) with experimental data on the cross sections are shown in Table 2. The $\text{T}^3(d, n)\text{He}^4$ and $\text{He}^3(d, p)\text{He}^4$ reactions show the effect of resonances and their reaction rates are considerably different from those obtained by Fermi and Turkevich.⁽²⁾

Table 2. Reaction rates $p(T)$

| Reactions | Reaction rates | |
|---|--|---|
| | Ours | Fermi-Turkevich |
| $p+n \rightarrow d+\gamma$ | $p_1 = 9.3 \times 10^{-20}$ | 6.6×10^{-20} |
| $d+d \rightarrow \text{He}^3+n$ | $p_2 = 9.00 \times 10^{-15} T_8^{-2/3} 10^{-5.18/T_8^{1/3}} *$ | $3.0 \times 10^{-15} T_8^{-2/3} 10^{-3.99/T_8^{1/3}}$ |
| $d+d \rightarrow \text{T}^3+p$ | $p_3 = 9.00 \times 10^{-15} T_8^{-2/3} 10^{-5.18/T_8^{1/3}}$ | $3.0 \times 10^{-15} T_8^{-2/3} 10^{-3.99/T_8^{1/3}}$ |
| $d+\text{T}^3 \rightarrow \text{He}^4+n$ | $p_4 = 5.48 \times 10^{-13} T_8^{-3/2} 10^{-3.92/T_8}$ | $5.0 \times 10^{-13} T_8^{-2/3} 10^{-4.24/T_8^{1/3}}$ |
| $d+\text{He}^3 \rightarrow \text{He}^4+p$ | $p_5 = 4.59 \times 10^{-13} T_8^{-3/2} 10^{-13.00/T_8}$ | $1.5 \times 10^{-12} T_8^{-2/3} 10^{-6.72/T_8^{1/3}}$ |

* $T_8 = T/10^8$

At the beginning of this stage, protons, neutrons, deuterons, tritons, He^3 and He^4 nuclei are found to be in statistical equilibrium as a whole. As the temperature falls, the inverse reactions of $\text{T}^3(d, n)\text{He}^4$ and $\text{He}^3(d, p)\text{He}^4$ become not so essential due to their large Q -values that the above equilibrium begins to break. However, a group of nuclei other than He^4 are closely in equilibrium among themselves due to their small Q -values. Therefore, using (3.4), we can express n_d , n_t and n_{He^3} in terms of n_n and n_p approximately as

$$n_d = \frac{1}{f_1} n_n n_p, \quad n_t = \frac{1}{f_1^2 f_2} n_n^2 n_p \quad \text{and} \quad n_{\text{He}^3} = \frac{1}{f_1^2 f_3} n_n n_p^2, \quad (3.9)$$

where

$$f_1 = 0.471 \left(\frac{2\pi HkT}{h^2} \right)^{3/2} \frac{1}{n} e^{-Q_1/kT} \quad (Q_1 = 2.23 \text{ Mev}),$$

$$f_2 = 3.46 e^{-Q_2/kT} \quad (Q_2 = 4.04 \text{ Mev}), \quad (3.10)$$

and

$$f_3 = 3.46 e^{-Q_3/kT} \quad (Q_3 = 3.26 \text{ Mev}).$$

Then, the time variation of n_α is given from (3.7) by

$$dn_\alpha/dt = n[p_4(T)\{n_d n_t - f_4 n_\alpha\} + p_5(T)\{n_d n_{\text{He}^3} - f_5 n_\alpha\}], \quad (3.11)$$

where

$$\begin{aligned} f_4 &= 5.61 e^{-Q_4/kT} \quad (Q_4 = 17.58 \text{ Mev}), \\ f_5 &= 5.61 e^{-Q_5/kT} \quad (Q_5 = 18.34 \text{ Mev}). \end{aligned} \quad (3.12)$$

The order estimation shows that this approximation is very good for $T \gtrsim 10^9$ °K. Taking an approximation that

$$\begin{aligned} n_d + n_t + 2n_{\text{He}^3} &\ll n_p + 2n_\alpha, \\ n_d + 2n_t + n_{\text{He}^3} &\ll n_n + 2n_\alpha, \end{aligned} \quad (3.13)$$

using (3.8) and (3.9) and replacing the variable t by $T_9 = T/10^9$, we have from (3.11)

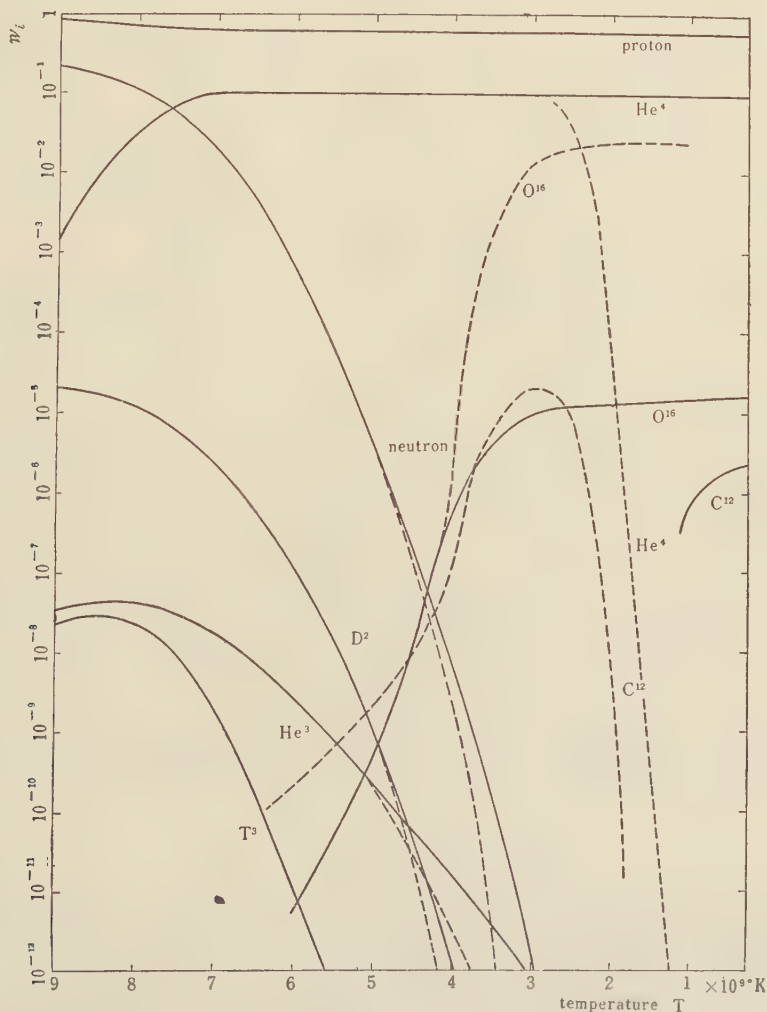
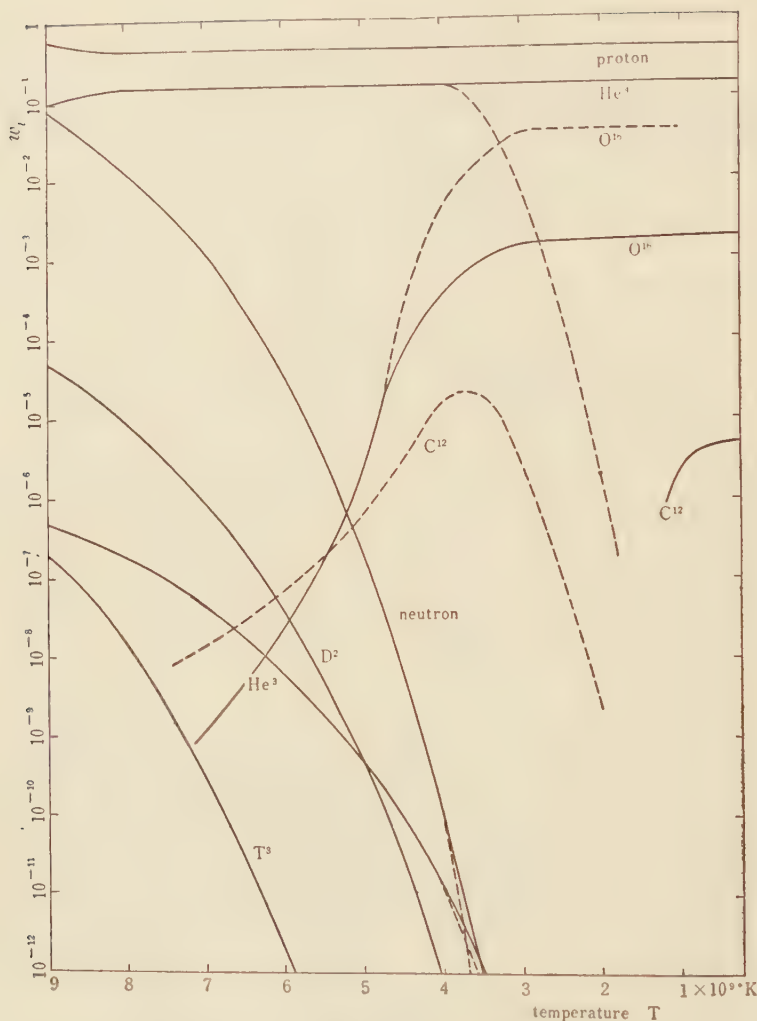


Fig. 2a. Temperature variation of the chemical abundances for $\rho_0 = 10^6 \text{ g/cm}^3$. Dotted curves represent their equilibrium values.

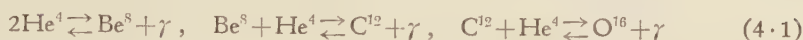
Fig. 2b. The same variation for $\rho_0=10^7$ g/cm³.

$$\begin{aligned}
 -\frac{dw_\alpha}{dT_0} = & 2.54 \times 10^{28} \left(\frac{\rho_0}{10^6} \right)^{1/2} T_0^{1/2} \left[p_4 \{ b - 2w_\alpha \} \left\{ \frac{1}{f_1^3 f_2} (b - 2w_\alpha)^2 (a - 2w_\alpha)^2 - f_4 w_\alpha \right\} \right. \\
 & \left. + p_5 \{ a - 2w_\alpha \} \left\{ \frac{1}{f_1^3 f_3} (b - 2w_\alpha)^2 (a - 2w_\alpha)^2 - f_5 w_\alpha \right\} \right]. \quad (3.14)
 \end{aligned}$$

From (3.8), (3.9), (3.13) and (3.14), all w_i 's are calculated, which are shown by the solid curves in Figs. 2a and 2b. As seen from Figs. 2a and 2b, amounts of neutrons, deuterons, tritons and He^3 are very small. Such small amount of neutrons is not sufficient to build up the heavy elements by neutron captures according to the scheme of the α - β - γ theory. After this stage, we must consider the reaction involving only protons and He^4 nuclei for the formation of nuclei heavier than He^4 .

§ 4. Formation of C^{12} and O^{16} nuclei

C^{12} and O^{16} nuclei are formed through a series of the nuclear reactions



from He^4 nuclei which have been formed in the stage considered in § 3. These reaction rates are calculated by Hayakawa et al.⁽⁶⁾ recently. Using their results, the changes in the relative abundances of C^{12} and O^{16} nuclei, w_C and w_O , are expressed as

$$-\frac{dw_C}{dT_9} = 2.54 \times 10^{28} \left(\frac{\rho_0}{10^6} \right)^{1/2} T_9^{1/2} [p_{Be \rightarrow C} \{w_{Be} w_\alpha - 4.36 \times 10^7 (\rho_0/10^6)^{-1} T_9^{-3/2} 10^{-37.18/T_9} w_C\} \\ - p_{C \rightarrow O} \{w_\alpha w_C - 5.20 \times 10^7 (10^6/\rho_0) T_9^{-3/2} 10^{-36.04/T_9} w_O\}], \quad (4.2a)$$

$$-\frac{dw_O}{dT_9} = 2.54 \times 10^{28} (\rho_0/10^6)^{1/2} T_9^{1/2} \\ \times [p_{C \rightarrow O} \{w_C w_\alpha - 5.20 \times 10^7 (10^6/\rho_0) T_9^{-3/2} 10^{-36.04/T_9} w_O\} - p_{O \rightarrow Ne} w_O w_\alpha], \quad (4.2b)$$

where

$$p_{Be \rightarrow C} = 1.92 \times 10^{-23} T_9^{-3/2} 10^{-1.390/T_9} \\ p_{C \rightarrow O} = 1.34 \times 10^{-14} T_9^{-2/3} 10^{-14.05/T_9^{1/3}}.$$

Since we can consider that $2He^4 \rightleftharpoons Be^8 + \gamma$ is always in equilibrium,⁽⁷⁾ w_{Be} is expressed in terms of w_α as

$$w_{Be} = 3.58 \times 10^{-8} T_9^{3/2} 10^{-0.484/T_9} (\rho_0/10^6) w_\alpha^2. \quad (4.3)$$

The second term in the square bracket of (4.2b) describes the decrease of O^{16} caused by the reaction $O^{16} + He^4 \rightarrow Ne^{20} + \gamma$. This term will be neglected in the calculations and its effect on the abundance of O^{16} will be discussed in § 5. We can consider the value of w_O in this approximation as the abundances of O^{16} together with all heavier nuclei. Then, if $w_O \gg w_{Ne}$, w_O gives the very abundance of O^{16} , and if $w_O \ll w_{Ne}$ it gives the abundances of Ne^{20} and heavier nuclei. Using an approximation $w_C < w_O$, we solve (4.2) for w_O and then w_C as the function of temperature by numerical integrations. The results are shown in Figs. 2a and 2b. Taking $\rho_0 = 10^6 \sim 10^7$ g/cm³, we can obtain the value of w_O which is close to the observed proton-oxygen abundance ratio, $1:10^{-4} \sim 10^{-3.9}$.*. However, the carbon-oxygen ratio obtained here is only $1/10 \sim 1/6$ of the observed ratio. A way to overcome the difficulty of low abundance of C^{12} nuclei will be suggested in § 5.

§ 5. Discussions

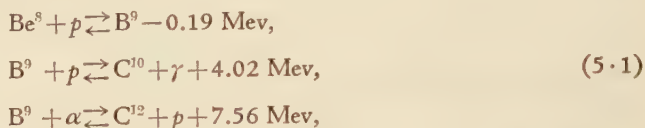
We will discuss the results of our calculations, the other routes for the formation

* According to the Taketani-Hatanaka-Obi theory,⁽⁵⁾ the abundances of elements formed in the pre-stellar stage should be compared with those in the stars of population II, in which the proton-oxygen ratio appears to be nearly $1:3 \times 10^{-4}$.

of C^{12} and O^{16} , the formation of Ne^{20} and heavier nuclei and the difficulties of low abundance of C^{12} shown in § 4.

(i) *Other routes to C^{12}*

We have considered only 3α -reactions in the formation of C^{12} , but it might be possible that C^{12} nuclei are formed through other reactions. Since at $T=4\sim 5\times 10^9$ °K, numbers of neutrons, deuterons, tritons and He^3 nuclei are found to be very small as shown in Figs. 2, we can neglect the possibility of forming C^{12} and heavier nuclei by the reactions involving nuclei other than protons and He^4 . Let us consider the routes of the formation from protons and He^4 only. Possible routes are



beside the one: $He^4 + p \rightleftharpoons Li^5 + \gamma$ and $Li^5 + \alpha \rightleftharpoons Be^8 + p$ producing again Be^8 which is in equilibrium with He^4 . At the high matter density under consideration, the reaction $Be^8(p, \gamma)B^9$ is in a state of equilibrium as in the case of $2He^4 \rightleftharpoons Be^8$. Then, we have

$$w_{B^9}/w_{Be^8} \sim 10^{-7} (\rho_0/10^6) T_9^{3/2} 10^{-0.93/T_9} w_p. \quad (5.2)$$

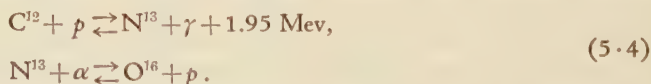
The ratio of reaction rates of $Be^8(\alpha, \gamma)C^{12}$ to $B^9(\alpha, p)C^{12}$ is given by

$$p_{Be \rightarrow C}/p_{B \rightarrow C} \simeq I_\gamma(Be^8 \rightarrow C^{12})/I_\alpha(B^9 \rightarrow C^{12}), \quad (5.3)$$

if the resonances are present in both reactions. From (5.2) and (5.3) it follows that the ratio of the number of C^{12} nuclei formed by $B^9(\alpha, p)C^{12}$ to the one by $Be^8(\alpha, \gamma)C^{12}$ is only $10^{-3} \sim 10^{-6}$. Therefore, we can neglect this route.

(ii) *Another route from C^{12} to O^{16}*

As another possible route from C^{12} to O^{16} , we can consider the following reactions.



At $T \lesssim 1.5 \times 10^9$ °K, the rate of O^{16} formation by the above route is found to be greater than that of the radiative α -capture. If we take account of this route, the w_C/w_O value calculated in § 4 is estimated to decrease remarkably.

(iii) *Formation of Ne^{20} nuclei*

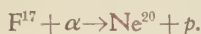
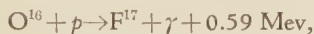
In calculating w_O in § 4, we neglected the reaction $O^{16}(\alpha, \gamma)Ne^{20}$. Since the photo-disintegration of Ne^{20} is negligible at $T \lesssim 3 \times 10^9$ °K, the relative abundance of Ne^{20} , w_{Ne} , can be calculated using the following equation

$$-\frac{dw_{Ne}}{dT_9} = 2.54 \times 10^{28} \left(\frac{\rho_0}{10^6} \right)^{1/2} T_9^{1/2} p_{O \rightarrow Ne} w_O w_\alpha. \quad (5.5)$$

At $T \gtrsim 0.6 \times 10^9$ °K we must use $p_{O \rightarrow Ne}$ which is calculated by Hayakawa et al.⁽⁶⁾ as the non-resonance case. But the spin and parity of the energy level of Ne^{20} is unknown and

Γ_τ in $O^{16}(\alpha, \gamma)Ne^{20}$ is uncertain. If we assume $\Gamma_\tau = 10^{-3}$ ev, we have $w_O \sim w_{Ne}$ as the final abundances. If Γ_τ is as large as 10^{-2} ev, we have $w_{Ne} \sim 10w_O$, which shows that we cannot neglect this reaction in the calculation of w_O in § 4. In this case O^{16} are used up in the formation of Ne^{20} and w_O will become as small as $1/10$ of w_{Ne} .

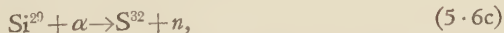
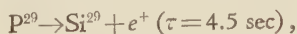
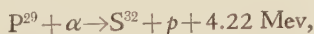
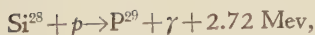
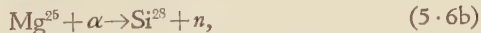
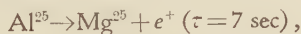
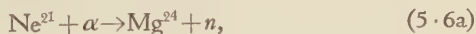
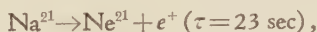
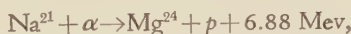
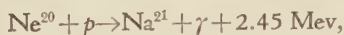
Let us consider another route of formation of Ne^{20} such as



This route will compete with the one, $O^{16}(\alpha, \gamma)Ne^{20}$, at $T \lesssim 10^9$ °K because of the low Q -value, 0.59 Mev, but will be neglected in the stage of Ne^{20} formation.

(iv) *Formation of nuclei heavier than Ne^{20}*

At $T \lesssim 10^9$ °K the number of remaining neutrons becomes so small that the formation of nuclei heavier than Ne^{20} seems to be possible only by p - and α -reactions. But in this stage the expansion of the universe is so slow that we cannot neglect the β -decays whose life-times are about $1 \sim 10$ sec. There exist the following reactions through which the heavier nuclei will be formed from proton, He^4 and Ne^{20} .



and so on. The (p, γ) reaction and the competition of β^+ -process with (α, p) reaction determine the number of neutrons formed by (α, n) reactions. It is desirable to calculate the amount of neutrons produced in these routes in order to see whether the formation of heavy nuclei is possible or not.

(v) *A way to reproduce C^{12}*

The calculated value of w_n in § 3 is much smaller than the observed one, which is $\sim 10^{-4}$.

The relative abundance of C^{12} nuclei, w_C , obtained in § 4 is also smaller than the

observed one. To solve this difficulty, we shall suggest the spallation, such as (p, d) (p, α) , of some O^{16} nuclei by non-thermal protons. Ginzburg,¹¹ and Hayakawa and Terashima¹² have suggested the existence of non-thermal protons accelerated in the expanding turbulent envelope of supernovae in order to explain the origin of cosmic rays. In our case also, it might be possible that turbulent motions accompanying the expansion of the universe will produce non-thermal protons in a similar way. If these relatively few protons could disintegrate some fraction of the O^{16} nuclei, say 1/10, the difficulty of the low abundance of C^{12} could be solved.

The authors wish to thank Professors M. Taketani and S. Hayakawa for valuable discussions and Professor H. Yukawa for continual encouragement of their work. They also wish to thank Mr. T. Tuzi for the preliminary calculations of the reaction rates used in § 3.

References

- 1) See, for example, R. A. Alpher and R. C. Herman, *Rev. Mod. Phys.* **22** (1950), 153.
- 2) E. Fermi and A. Turkevich, see ref. (1).
- 3) F. Hoyle, *M. N. R. A. S.* **106** (1946), 343; *Ap. J. Suppl.* **1** (1954), 121.
- 4) W. A. Fowler and J. L. Greenstein, *Proc. Nat. Acad. Sci.* **42** (1956), 173.
- 5) M. Taketani, T. Hatanaka and S. Obi, *Prog. Theor. Phys.* **15** (1956), 89.
- 6) S. Hayakawa, C. Hayashi, M. Imoto and K. Kikuchi, *Prog. Theor. Phys.* **16** (1956) (in press).
- 7) C. Hayashi, *Prog. Theor. Phys.* **5** (1950), 224.
- 8) R. A. Alpher, J. W. Follin, Jr. and R. C. Herman, *Phys. Rev.* **92** (1953), 1347.
- 9) H. E. Suess and H. C. Urey, *Rev. Mod. Phys.* **28** (1956) 53.
- 10) W. A. Fowler, G. R. Burbidge and E. M. Burbidge, *Ap. J.* **122** (1955), 271; *Ap. J. Suppl.* **17** (1955), 167.
- 11) V. L. Ginzburg, *Dokl. Akad. Nauk U.S.S.R.* **92** (1953), 727; *Usp. Fiz. Nauk.* **51** (1953), 343; *Izv. Acad. Nauk U.S.S.R.* **20** (1955), 5.
- 12) S. Hayakawa and Y. Terashima, private communication.

The Composite Model for New Unstable Particles, I

Sho TANAKA

Department of Physics, Rikkyo University, Tokyo

(Received September 4, 1956)

Models for K -mesons are mainly discussed from the view-point of regarding new unstable particles as composite particles. To interpret the degeneracy of masses of K -mesons, we assume the existence of two kinds of Λ -particles together with nucleon as fundamental particles and consider several models for K -mesons

§ 1. Introduction

Recently, Sakata¹⁾ has proposed a "compound" hypothesis for new unstable particles to account for the Nishijima²⁾ and Gell-Mann's³⁾ rule (N-G's rule). In his model new particles are considered to be composed of a nucleon, a Λ -particle, and their antiparticles. The basic idea characteristic to his theory is that the η -charge or the strangeness which is assigned to a new particle and plays a fundamental role in the N-G's theory, is identified as the number of Λ -particles involved in the model of the new particle, and that the conservation of the η -charge in the strong interaction is attributed to the conservation of the number of Λ -particles. We may, therefore, consider that the mysterious quantity η introduced phenomenologically in the N-G's theory is confined to the entity of Λ -particle itself in the Sakata's theory.

Furthermore, the requirement for the iso-invariance of all the strong interactions in the N-G's theory is reduced to that for the invariance of the basic interactions of fundamental particles. Since the effective interactions of composite particles come from the basic interaction, the invariance of the former follows automatically from that of the latter.

Therefore, we may consider that the concepts of the η -charge and the iso-spin characteristic to the N-G's theory are reduced to the problem of the structure of the basic interaction among the fundamental particles, that is, the problem of the conservation of the number of Λ -particles and of the choice of the suitable arrangement of neutrons and protons in the basic interactions.

In the present paper, focusing our attention on the degeneracy of masses of K -mesons we reexamine, in particular, the models for K -mesons and attempt the modification of the Sakata's model. Although there are many interesting phenomena with regard to the new particles, it may be worth-while to remark the fact that several K -events show different decay modes, nevertheless they have no appreciable mass difference. At present it is not clear whether only one kind of particle is responsible for these K -events. However, we have some evidence⁴⁾ which seems favourable for the existence of two kinds of K -mesons,

i.e., θ and τ , though they have nearly equal masses. So far several attempts have been made to interpret those K -mesons from the theoretical view-point.^{5), 6)} Now, according to the Sakata's model, these two kinds of K -meson, θ and τ , are interpreted as two bound states of a nucleon and a \bar{A} -particle with different internal structures, e.g., 1S_0 state and 3P_0 state. However, it seems very difficult, from such a standpoint, to account for the fact that the two particles are considerably stable and have nearly degenerated masses. Practically, one can equate these two energy levels by adjusting the several parameters, e.g., the strength of coupling or the cut-off momentum, etc., but such an agreement is, we should consider, extremely accidental.

Thus, it seems quite reasonable to regard this degeneracy of K -mesons as coming from a more fundamental basis. Therefore we shall assume, as a possibility to overcome the difficulty, that there are two kinds of A -particles, i.e., A and A' , and they have equal masses and equal spins $1/2$ and are distinguished only by their opposite space parity. From this point of view, we shall discuss in the next section, several possible models for K -mesons, taking into account the detailed form of interactions among nucleon, A , and A' -particles.

§ 2. Models for K -mesons

As mentioned in the previous section, the Sakata's model for new unstable particles has succeeded in accounting for N-G's rule. This model is illustrated in Table I, where the symbols N , A and \bar{N} , \bar{A} stand for nucleon, A -particle, and their antiparticles.

Table I

| Name | Sakata's model (spin, parity) | Modified model |
|-------------------|------------------------------------|---|
| $\pi^{\pm}, 0$ | $N\bar{N} \quad (0, -)$ | $N\bar{N}$ |
| $\theta^{+}, 0$ | $N\bar{A} \quad (0, +)$ | $N\bar{A}$, or $N\bar{A}_1, N\bar{A}_2$ |
| $\tau^{+}, 0$ | $N\bar{A}' \quad (0, -)$ | $N\bar{A}'$, or $N\bar{A}'_1, N\bar{A}'_2$ |
| $\Sigma^{\pm}, 0$ | $AN\bar{N} \quad (\frac{1}{2}, -)$ | $ANN, A'N\bar{N}$ |
| $\Xi^{0,-}$ | $AA\bar{N} \quad (\frac{1}{2}, -)$ | $AA\bar{N}, AA'\bar{N}, A'A\bar{N}$ |

As the fundamental interactions sufficient to make possible these compound models, the following Fermi-interactions can be considered :

(I)
$$\sum_{i=1}^5 G_i^{NN} N \tau_i \varrho_i N \bar{N} \tau_i \varrho_i N + \sum_{i=1}^5 G_i^{A'NN} \bar{N} \tau_i \varrho_i N \bar{N} \tau_i \varrho_i N, \tag{2.1}$$

$$\sum_{i=1}^5 G_i^{\Lambda N} \bar{N} \tau_i \varrho_i N \bar{A} \varrho_i A, \tag{2.2}$$

$$\sum_{i=1}^5 G_i^{\Lambda A} \bar{A} \varrho_i A \bar{A} \varrho_i A, \tag{2.3}$$

where
$$N = \begin{pmatrix} P \\ N \end{pmatrix}.$$

However, as pointed out in the preceding section, Sakata's model implies some difficulty concerning the mass degeneracy of K -mesons, and we proposed the existence of two kinds of Λ -particles, Λ and Λ' , to get rid of this difficulty.

In this case, several models for K -mesons may be possible, corresponding to the form of interactions among nucleon, Λ -, and Λ' -particles. In the following, we shall confine ourselves to a rather simple interaction form for the sake of simplicity. i) Let us consider first a case in which each number of Λ or Λ' is conserved independently. In addition, if we remember the conservation of the electric charge and the number of heavy particles, one can easily find that, in this case, such interaction necessarily leads to the conservation of each number of fundamental particles, N , P , Λ , or Λ' . As an interaction of this kind, the following one may be conceivable in place of (I)

$$\sum_{i=1}^5 G_i^{NN} \bar{N} \vec{\tau} \cdot \underline{Q}_i N \bar{N} \cdot \vec{\tau} \cdot \underline{Q}_i N + \sum_{i=1}^5 G_i'^{NN} \bar{N} \tau_i \cdot \underline{Q}_i N \bar{N} \tau_i \cdot \underline{Q}_i N, \quad (2.1)'$$

$$(I)' \quad \sum_{i=1}^5 G_i^{\Lambda N} \bar{N} \tau_i \cdot \underline{Q}_i N \bar{\Lambda} \tau_i' \cdot \underline{Q}_i \Lambda, \quad (2.4)'$$

$$\sum_{i=1}^5 G_i^{\Lambda \Lambda} \bar{\Lambda} \vec{\tau} \cdot \underline{Q}_i \Lambda \cdot \bar{\Lambda} \vec{\tau} \cdot \underline{Q}_i \Lambda + \sum_{i=1}^5 G_i'^{\Lambda \Lambda} \bar{\Lambda} \tau_i' \cdot \underline{Q}_i \Lambda \bar{\Lambda} \tau_i' \cdot \underline{Q}_i \Lambda, \quad (2.5)'$$

where $\Lambda = \begin{pmatrix} \Lambda \\ \Lambda' \end{pmatrix}$, and τ_μ and τ_μ' are Pauli's 2×2 matrices operating on N and Λ , respectively. Hereafter, we shall denote by Λ and Λ' those particles which have the opposite and the same parities relative to nucleon, respectively.

Now, introducing the fields, $\Lambda_1 = (\Lambda + \Lambda')/\sqrt{2}$ and $\Lambda_2 = (\Lambda - \Lambda')/\sqrt{2}$ in place of Λ and Λ' , we can rewrite (2.4)' in the form:

$$\sum_{i=1}^5 G_i^{\Lambda N} \bar{N} \tau_i \cdot \underline{Q}_i N (\bar{\Lambda}_1 \underline{Q}_i \Lambda_1 + \bar{\Lambda}_2 \underline{Q}_i \Lambda_2). \quad (2.4)''$$

Here, both Λ_1 and Λ_2 are obviously indefinite as to space parity, whereas they satisfy the same equation of motion as Λ or Λ' .

These interactions, (2.4)' or (2.4)'', give rise to stationary states of N and $\tilde{\Lambda}$ or $\tilde{\Lambda}'$ system, as listed in Table II.

Table II

| Name | model | state | parity |
|--|----------------------|--------|--------|
| $K_e^{+,0}$ | $N\tilde{\Lambda}$ | $1S_0$ | even |
| $K_0^{+,0}$ | $N\tilde{\Lambda}'$ | $1S_0$ | odd |
| $K_1^{+,0} = (K_e^{+,0} + K_0^{+,0})/\sqrt{2}$ | $N\tilde{\Lambda}_1$ | $1S^0$ | mixed |
| $K_2^{+,0} = (K_e^{+,0} - K_0^{+,0})/\sqrt{2}$ | $N\tilde{\Lambda}_2$ | $1S_0$ | mixed |

It is clear that K_e , K_0 , K_1 , and K_2 have the same energies in the center of mass

system, that is, the same masses. We can find immediately from the space parity consideration that K_c and K_0 may have decay modes of θ - and τ - events, respectively. On the other hand, K_1 (and also K_2) can decay either into 2π 's through K_c component or 3π 's through K_0 -component. In our case, it is an interesting question which of K_c , K_0 , and $K_{1,2}$ is to be identified as the actual θ - and τ - mesons discovered in nature.

Let us consider the mechanism of production of K -mesons at the collision of a π^- and a proton. Considering $(2.4)'$ and $(2.4)''$, we find that the following processes may occur;

$$\pi^- + p \rightarrow \begin{cases} K_c + \Lambda \\ K_0 + \Lambda' \\ K_1 + \Lambda_1 \\ K_2 + \Lambda_2 \end{cases} \quad (2.4)$$

Since the decay constants of Λ and Λ' may generally differ from each other as will be seen in the next section, K_c and K_0 may also be expected to have different lifetimes. On the other hand, K_1 (and also K_2), which has two alternative decay modes, must be considered as "particle mixture".⁷⁾

ii) Secondly, we shall examine another interaction as follows:

$$\begin{aligned} & \sum_{i=1}^5 G_i^{\Lambda N} \bar{N} \tau_i \Omega_i N (\bar{\Lambda} \Omega_i \gamma_5 \Lambda' + \bar{\Lambda}' \Omega_i \gamma_5 \Lambda) \\ &= \sum_{i=1}^5 G_i^{\Lambda N} \bar{N} \tau_i \Omega_i N (\bar{\Lambda}_1 \Omega_i \gamma_5 \Lambda_1 - \bar{\Lambda}_2 \Omega_i \gamma_5 \Lambda_2). \end{aligned} \quad (2.5)$$

Although, in this interaction, Λ and Λ' are not independently conserved, they appear in a symmetric manner. In the case of this interaction, the following bound states in Table III are conceivable.

Table III

| Name | model | parity |
|---------|---|--------|
| K_c' | $[N\tilde{\Lambda}(^1S_0) + N\tilde{\Lambda}'(^3P_0)]/\sqrt{2}$ | even |
| K_o' | $[N\tilde{\Lambda}(^3P_0) + N\tilde{\Lambda}'(^1S_0)]/\sqrt{2}$ | odd |
| K_c'' | $[N\tilde{\Lambda}(^1S_0) - N\tilde{\Lambda}'(^3P_0)]/\sqrt{2}$ | even |
| K_o'' | $[N\tilde{\Lambda}(^3P_0) - N\tilde{\Lambda}'(^1S_0)]/\sqrt{2}$ | odd |

Looking at the symmetrical appearance of Λ and Λ' in (2.5), we find immediately that K_c' and K_o' belong to the same energy level and also K_c'' and K_o'' belong to the other same energy level. Therefore, in this model, we can identify the two states in a lower energy level or their-linear combinations to the θ -meson and the τ -meson, because the states in a higher energy level may decay into the states in a lower energy level in a very short time.

Finally we shall turn our attention to the strange case where, in the interaction (2.5), γ_5 is replaced by 1. Of course, in this case, the interaction Hamiltonian becomes pseudoscalar quantity with respect to the space reflection. However, if such an unfamiliar interaction is permissible without causing any physical difficulty, we find that either of the following states,

$$K_1' = N\tilde{A}_1(^1S_0) = (1/\sqrt{2})[N\tilde{A}(^1S_0) + N\tilde{A}'(^1S_0)]$$

and

$$K_2' = N\tilde{A}_2(^1S_0) = (1/\sqrt{2})[N\tilde{A}(^1S_0) - N\tilde{A}'(^1S_0)], \quad (2.6)$$

can be a bound state, taking into consideration the second expression of (2.5). Since the first and the second components of $K'_{1,2}$ have even and odd space parities, respectively, K_1' or K_2' becomes a mixed state of space parity, just as K_1 and K_2 . Thus, $K'_{1,2}$ have a similar property to K_1 and K_2 discussed in i), but in $K'_{1,2}$ the first and the second components are dynamically combined by the strong interaction

$$N\tilde{A} \rightleftharpoons N\tilde{A}' \quad (2.7)$$

in contrast to the case of $K_{1,2}$.*

So far we have discussed only the K -meson models. Obviously the introduction of two kinds of A -particles necessarily causes the modification of the models for hyperons as shown in the third column in Table I, although here we shall not study in detail.

§ 3. Concluding remarks

Now, it should be noted that our attempt discussed above seems to be related closely with Lee and Yang's⁽¹⁾ theory concerning K -mesons. Indeed, the operation of parity conjugation introduced by Lee and Yang corresponds to the exchange of A and A' in our case. With respect to the differences between both theories, one may tentatively summarize them as follows. (i) In our theory the transformation property of new particles under the parity conjugation, that is, the exchange of a A and a A' -particles involved in the model of new particles, may be uniquely determined by the internal structure of composite particles, without any *ad hoc* assumption, contrary to L-Y's theory. (ii) As shown in (2.4), it is uniquely determined which kind of K -mesons is produced in association with A or A' at the collision of a π^- meson and a proton, while in L-Y's theory the processes

$$\pi^- + P \rightarrow A + \tau^0, \quad \pi^- + P \rightarrow A' + \theta^0 \quad (3.1)$$

may be possible as well as

$$\pi^- + P \rightarrow A + \theta^0, \quad \pi^- + P \rightarrow A' + \tau^0, \quad (3.2)$$

* If we do not replace γ_5 with 1 in the interaction (2.5), there does not exist a bound state in a similar form to that of $K'_{1,2}$, because the expectation value of the interaction Hamiltonian (2.5) in the state of K_1' or K_2' vanishes.

even though production rates of both processes (3·1) and (3·2) are generally different.

Further, we have to notice that other models of this kind for new particles may also be possible. Actually, Goldhaber⁸⁾ has presented a model, in which a nucleon and θ -mesons are considered as fundamental particles and θ -mesons play just the same role as a M -particle in the Sakata's model. However, at present, it seems premature to conclude which one is superior to the other, though some differences of both models will be illustrated in a forthcoming paper.

Finally, the author would like to express his sincere thanks to Professor S. Sakata for his kind interest and valuable discussions.

References

- 1) S. Sakata, *Prog. Theor. Phys.* **16** (1956), 686.
- 2) K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285.
- 3) M. Gell-Mann, *Phys. Rev.* **92** (1953), 833.
- 4) J. Orear, G. Harris, and S. Taylor, *Phys. Rev.* **102** (1956), 1676.
- 5) T. D. Lee and J. Orear, *Phys. Rev.* **100** (1955), 932.
- 6) T. D. Lee and C. N. Yang, *Phys. Rev.* **102** (1956), 290.
- 7) M. Gell-Mann and A. Pais, *Phys. Rev.* **97** (1955), 1387.
A. Pais and O. Piccioni, *Phys. Rev.* **100** (1955), 1487.
- 8) M. Goldhaber, *Phys. Rev.* **101** (1956), 433.

The Composite Model for New Unstable Particles, II

Sho TANAKA

Department of Physics, Rikkyo University, Tokyo

(Received September 4, 1956)

Decays of the composite particles are discussed from the same point of view as discussed in the preceding article, I. The composite particles, which are composed of fundamental particles by strong Fermi-interactions, are assumed to disintegrate through the several universal weak Fermi-couplings. It is examined from the standpoint of the universal weak Fermi-couplings whether the conditions required for the internal wave functions of composite π and K mesons from the experimental results, i.e., observed lifetimes or masses of new particles, are mutually consistent.

§ 1. Introduction

The composite π and K mesons combined through the strong Fermi-couplings (I) or (I)' in the previous paper¹⁾ (to be referred to as I in what follows) are considered to disintegrate via several universal weak Fermi-couplings, as it will be shown later. Although decay constants for these disintegrations may be calculated only when the internal wave functions of composite particles are obtained, we have so far no satisfactory method to treat the problem of such a bound state.²⁾³⁾⁴⁾ However, on the contrary, we may expect that some information on the behavior of the wave functions of the composite particles can be obtained from the lifetimes or masses of new particles without any knowledge of the forces which make possible to composite particles.⁵⁾ In the present paper, we wish to investigate whether the conditions for the internal wave functions of composite π - and K -mesons required from the experimental results, observed lifetimes or masses of new particles, are mutually consistent.

§ 2. Decays of new particles

i) $\pi^+ \rightarrow \mu^+ + \nu$

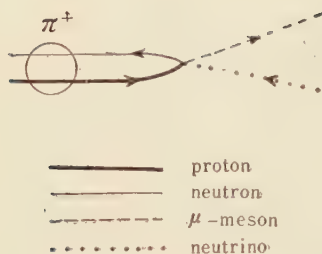
π^+ -meson may decay into $\mu^+ + \nu$ via the well-known μ -capture interaction,

$$H_\mu = \sum_{i=1}^5 g_i \bar{N} \mathcal{O}_i P \bar{\mu} \mathcal{O}_i \nu. \quad (2.1)$$

This process is illustrated by the diagram in Fig. 1. The matrix element in the first order for this process is

$$\langle \mu^+, \nu | \bar{H}_\mu | \pi^+ \rangle, \quad (2.2)$$

where \bar{H}_μ is $\int d^3 X H_\mu$ and $|\pi^+ \rangle$ stands for the state of a composite π^+ -meson ($P\tilde{N}$: 1S_0 -state) at rest. We shall assume the following two body approximation form

Fig. 1. $\pi^+ \rightarrow \mu^+ + \nu$

$$|\pi^+\rangle = \sum_{l, \alpha, \beta} a^{r*}(l) b^{s*}(-l) u_{\alpha}^{r*}(l) \varphi_{\alpha\beta}^{\pi^+}(\mathbf{r}) u_{\beta}^{s*}(-l) \times e^{-i\mathbf{l}\cdot\mathbf{r}} \int d^3r |0\rangle, \quad (2.3)$$

where $|0\rangle$ means a vacuum state and $a^r(l)$ and $b^r(l)$ are the annihilation operators of a proton and an antineutron with the momentum l and the spin polarization r , respectively. $\varphi_{\alpha\beta}^{\pi^+}(\mathbf{r})$'s represent the internal wave function of a π^+ -meson at rest with 16-components, where \mathbf{r} denotes the relative coordinate between a proton and an antineutron. Further, $u_{\alpha}^r(l)$'s represent a

constant spinor of a nucleon with the momentum l and the spin polarization r .

One can develop $\varphi_{\alpha\beta}^{\pi^+}(\mathbf{r})$ into tensor components⁵⁾ as

$$\varphi_{\alpha\beta}^{\pi^+}(\mathbf{r}) = \sum_{i=1}^5 \lambda_i^{\pi}(\mathbf{r}) \Omega_i C, \quad (2.4)$$

where C is a charge conjugation operator defined by $C \gamma_{\mu}^T C^{-1} = -\gamma_{\mu}$, and $\lambda_i^{\pi}(\mathbf{r})$'s consist of scalar $\lambda^{\pi}(\mathbf{r})$; vector $\mathbf{a}^{\pi}(\mathbf{r})$, $\mathbf{a}_i^{\pi}(\mathbf{r})$; tensor $\mathbf{H}^{\pi}(\mathbf{r})$, $\mathbf{E}^{\pi}(\mathbf{r})$; axial vector $\mathbf{\lambda}^{\pi}(\mathbf{r})$, $\lambda_i^{\pi}(\mathbf{r})$; and pseudoscalar $P^{\pi}(\mathbf{r})$.

Then, we obtain for the lifetime of a π^+ -meson at rest

$$\tau^{-1} = (16/\pi) q_{\mu}^0 q_{\nu}^0 q m_{\pi}^{-1} \times [|g_4 \lambda_4^{\pi}(0)|^2 (1 - q/q_{\mu}^0) + |g_5 P^{\pi}(0)|^2 (1 + q/q_{\mu}^0) + 2\Re_i g_4 g_5^* \lambda_4^{\pi}(0) P^{\pi*}(0) m_{\mu}/q_{\mu}^0], \quad (2.5)^{*)},^{**})$$

where

q_{μ}^0, q_{ν}^0 : energies of μ and ν ,

q : momentum of both particles,

m_{π}, m_{μ} : masses of a π - and a μ -mesons,

and the symbol \Re means to take out the real part of the subsequent term.

Inserting the observed value 3×10^{-8} sec for τ in (2.5), we obtain the following results:

a) In the case of $g_5=0$ and $g_4=g$,

$$|\lambda_4^{\pi}(0)| = 0.75 \times 10^{-2} M^{3/2}, \quad (2.6)$$

b) In the case of $g_4=0$ and $g_5=g$,

$$|P^{\pi}(0)| = 0.60 \times 10^{-2} M^{3/2}. \quad (2.7)$$

*) In the following, we shall use the relativistic unit,

$$\hbar = c = 1.$$

**) The process $\pi \rightarrow e + \nu$ may be obtained only if the quantities with respect to the μ -meson are replaced by ones with respect to the electron. From (2.5), one finds that the unwanted process $\pi \rightarrow e + \nu$ can be excluded if the usual Fermi-couplings of β -decay consist of only scalar and tensor interactions.⁶⁾ See also 5).

g denotes the well-known universal Fermi-coupling constant $\sim 2 \times 10^{-10}$ erg $\text{cm}^3 = 1.5 \times 10^{-5} M^{-2}$, and M represents the nucleon mass.

$$\text{ii) } K_{u2}^+ \rightarrow \mu^+ + \nu$$

$K_e(P\tilde{A}: {}^1S_0\text{-state})$ and also $K_0(P\tilde{A}': {}^1S_0\text{-state})$ (or K_0 - and K_e -components of K_{132}), which are discussed in I, are assumed to decay into $\mu^+ + \nu$ through the weak Fermi-couplings of the form

$$H_\mu^\Lambda = \sum_{i=1}^5 g_i \bar{A} \gamma_5 \Omega_i P \bar{\nu} \Omega_i \mu, \quad (2.8)$$

$$H_\mu^{\Lambda'} = \sum_{i=1}^5 g_i' \bar{A}' \Omega_i P \bar{\nu} \Omega_i \mu. \quad (2.9)$$

This process is illustrated by the diagram in Fig. 2.

In the same way as in π -decay, we obtain for the lifetimes of τ_e and τ_0 of K_e and K_0 ,

$$\begin{aligned} \tau_{\{e\}}^{-1} &= (16/\pi) q_\mu^0 q_\nu^0 q m_K^{-1} \\ &\times \left[\left| \left\{ -\frac{g_2}{g_4'} \right\} \lambda_4^0(0) \right|^2 (1 - q/q_\mu^0) + \left| \left\{ \frac{g_1}{g_5'} \right\} I^0(0) \right|^2 (1 + q/q_\mu^0) \right. \\ &\quad \left. + 2\Re \left\{ -\frac{g_1}{g_5'} \frac{g_2^*}{g_4'^*} \right\} I^0(0) \lambda_4^{0*}(0) m_\mu/q_\mu^0 \right], \end{aligned} \quad (2.10)$$



where the first and the second rows in the brackets $\{ \}$ refer to K_e - and K_0 -mesons, respectively, and m_K represents the mass of the K -meson.

Inserting the observed value 10^{-8} sec for τ_e and τ_0 into (2.10), we obtain the following results:

a) In the case of $g_2 = g_4' = 0$ and $g_1 = g_5' = g$,

$$|I^0(0)| = 1.0 \times 10^{-3} M^{3/2}. \quad (2.11)$$

b) In the case of $g_1 = g_5' = 0$ and $g_2 = g_4' = g$,

$$|\lambda_4^0(0)| = 4.7 \times 10^{-3} M^{3/2}. \quad (2.12)$$

iii) $K^+ \rightarrow \mu^+ + \nu + \pi^0$

K_e and K_0 may also decay into $\mu^+ + \nu + \pi^0$ through the interactions (2.8) and (2.9). This process is illustrated by the diagram in Fig. 3.

The lifetime of this process is given approximately as follows:

$$\tau_{\{e\}}^{-1} = \frac{m_K^4}{8\pi^3} \int_{m_\mu}^{m_K/2} dx \sqrt{x^2 - m_\mu^2} (m_K/2 - x)^2 (m_K - x)^{-4}$$

$$\times \left[x \left(\left| \left\{ -\frac{g_4}{g_2'} \right\} \mathfrak{F} \right|^2 + \left| \left\{ \frac{g_5}{g_1'} \right\} \mathfrak{G} \right|^2 \right) + 2\Re \left\{ -\frac{g_4}{g_2'} \frac{g_5^*}{g_1'^*} \right\} \mathfrak{F} \mathfrak{G}^* \right]. \quad (2.13)$$

Fig. 2. $K^+ \rightarrow \mu^+ + \nu$

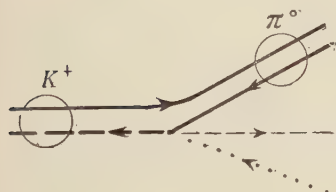


Fig. 3. $K^+ \rightarrow \mu^+ + \nu + \pi^0$

In the above expression,

$$\tilde{F} = 4 \{ d^3 r [I^{\pi*}(r) I^{\pi^0}(r) + \lambda_4^{\pi*}(r) \lambda_4^{\pi^0}(r) + H^{\pi*}(r) \cdot H^{\pi^0}(r)] \quad (2.14)$$

and

$$\mathbb{G} = 4 \{ d^3 r [I^{\pi*}(r) \lambda_4^{\pi^0}(r) - \lambda_4^{\pi*}(r) I^{\pi^0}(r)] \}. \quad (2.15)$$

\tilde{F} cannot exceed 1 due to the normalization of the wave functions of π - and K -mesons. That is, in this case, the relation between the lifetime and the wave functions of π - and K -mesons is not so simple as in (2.5) and (2.10).

Inserting the observed value 10^{-8} sec for τ_e and τ_0 into (2.13), we obtain the following results:

a) In the case of $g_5 = g_1' = 0$ and $g_4 = g_2' = g$,

$$|\tilde{F}| = 1.1. \quad (2.16)$$

b) In the case of $g_4 = g_2' = 0$ and $g_5 = g_1' = g$,

$$|\mathbb{G}| = 1.1. \quad (2.17)$$

$$\text{iv) } \theta^0 \rightarrow \begin{cases} \pi^+ + \pi^- \\ \pi^0 + \pi^0 \end{cases}, \quad K^+ \rightarrow \pi^+ + \pi^0$$

The disintegration of K -meson into two pions may be explained as the decay of $K_e(N\tilde{A}: {}^1S_0\text{-state})$ (or K_e -component of $K_{1,2}$) through the interaction of the following form

$$\begin{aligned} H_N^A &= \sum_{i=1}^0 g_i \bar{A} \gamma_5 Q_i N \bar{N} Q_i N + \sum_{i=1}^b \bar{g}_i \bar{A} \gamma_5 Q_i P \bar{P} Q_i N \\ &= - \sum_{i=1}^5 \tilde{g}_i \bar{A} \gamma_5 Q_i N \bar{N} Q_i N - \sum_{i=1}^5 \tilde{\bar{g}}_i \bar{A} \gamma_5 Q_i N \bar{P} Q_i P. \end{aligned} \quad (2.18)$$

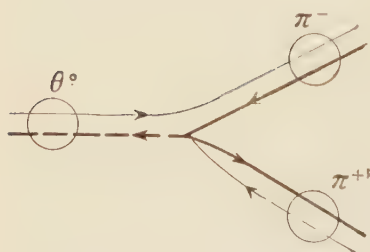


Fig. 4. $\theta^0 \rightarrow \pi^+ + \pi^-$

In the above expression, \tilde{g}_i and $\tilde{\bar{g}}_i$ are quantities which are suitable linear combinations of g_i and \bar{g}_i , respectively. These linear combinations are determined by Fierz's identity⁷⁾ of γ -matrix. One of three processes, for example, the process, $\theta^0 \rightarrow \pi^+ + \pi^-$, is illustrated by the diagram in Fig. 4.

The lifetimes are given as follows:

$$\begin{aligned} \tau^{-1} \left\{ \begin{array}{l} \theta^0 \rightarrow \pi^+ + \pi^- \\ \theta^0 \rightarrow \pi^0 + \pi^0 \\ K^+ \rightarrow \pi^+ + \pi^0 \end{array} \right\} &= \\ &= \frac{8}{\pi} m_K \sqrt{(m_K/2)^2 - m_\pi^2} \end{aligned}$$

$$\times \left[\left\{ \frac{\bar{g}_5}{g_5 - (\bar{g}_5 - \tilde{g}_5)} \right\} \Gamma^{\pi^*}(0) \mathfrak{G} + \left\{ \frac{\bar{g}_4}{g_4 - (\bar{g}_4 - \tilde{g}_4)} \right\} \Gamma^{\pi^*}(0) \mathfrak{F} \right] \quad (2.19)$$

Inserting the observed value 10^{-10} sec for $\tau_{\pi^0 \rightarrow \pi^+ + \pi^-}$ into (2.19), we obtain the following results:

a) In the case of $\bar{g}_5=0$ and $\bar{g}_4=g$,

$$|\lambda_4^{\pi^*}(0) \mathfrak{F}| = 1.1 \times 10^{-2} M^{3/2}. \quad (2.20)$$

b) In the case of $\bar{g}_4=0$ and $\bar{g}_5=g$,

$$|\Gamma^{\pi^*}(0) \mathfrak{G}| = 1.1 \times 10^{-2} M^{3/2}. \quad (2.21)$$

Substituting the values of \mathfrak{F} and \mathfrak{G} , i.e., (2.16) and (2.17) into the above equations, we obtain

$$a) \quad |\lambda_4^{\pi}(0)| = 1.0 \times 10^{-2} M^{3/2}, \quad (2.22)$$

$$b) \quad |\Gamma^{\pi}(0)| = 1.0 \times 10^{-2} M^{3/2}. \quad (2.23)$$

Combining the observed value 10^{-8} sec for $\tau_{K^+ \rightarrow \pi^+ + \pi^0}$ and the values $|\lambda_4^{\pi^*}(0) \mathfrak{F}|$ and $|\Gamma^{\pi^*}(0) \mathfrak{G}|$ calculated by (2.20) and (2.21) in the expression for $\tau_{K^+ \rightarrow \pi^+ + \pi^0}$ shown by (2.19), we obtain the following relation,

$$\begin{aligned} a) \quad & (g_5 - \bar{g}_5) - (\tilde{g}_5 - \tilde{\bar{g}}_5) = 0.14g, \\ b) \quad & (g_4 - \bar{g}_4) - (\tilde{g}_4 - \tilde{\bar{g}}_4) = 0.14g, \end{aligned} \quad (2.24)$$

which will be discussed in the next section.

$$v) \quad A \text{ or } A' \rightarrow \begin{cases} P + \pi^- \\ N + \pi^0 \end{cases}$$

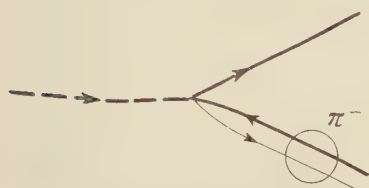


Fig. 5. $A \rightarrow P + \pi^-$

A -particle may disintegrate into $P + \pi^-$ or $N + \pi^0$ through H_N^A given by (2.18) and also A' -particle through an interaction $H_N^{A'}$ similar to H_N^A , i.e.,

$$\begin{aligned} H_N^{A'} &= \sum_{i=1}^5 g_i' \bar{A}' \mathcal{Q}_i N \bar{N} \mathcal{Q}_i N + \sum_{i=1}^5 \bar{g}_i' \bar{A}' \mathcal{Q}_i P \bar{P} \mathcal{Q}_i N \\ &= - \sum_{i=1}^5 \tilde{g}_i' \bar{A}' \mathcal{Q}_i N \bar{N} \mathcal{Q}_i N - \sum_{i=1}^5 \tilde{\bar{g}}_i' \bar{A}' \mathcal{Q}_i N \bar{P} \mathcal{Q}_i P. \end{aligned} \quad (2.18)'$$

One of these two processes, $A \rightarrow P + \pi^-$, for example, is illustrated by the diagram in Fig. 5.

Then, lifetimes of both particles are given, respectively, as follows:

$$\begin{aligned}
& \tau^{-1} \left\{ \begin{array}{l} A \rightarrow P + \pi^- \\ A \rightarrow N + \pi^0 \end{array} \right\} = \\
& = \frac{16}{\pi} q_P^0 q_\pi^0 q m_\Lambda^{-1} \\
& \times \left[\left| \left\{ \frac{\bar{g}_1}{\tilde{g}_1 - (\tilde{g}_1 - g_1)} \right\} \Gamma^\pi(0) \right|^2 + \left| \left\{ \frac{\bar{g}_2}{\tilde{g}_2 - (\tilde{g}_2 - g_2)} \right\} \lambda_1^\pi(0) \right|^2 \right. \\
& \left. + 2\Re i \left\{ \frac{\bar{g}_1^* \tilde{g}_2}{\tilde{g}_1^* - (\tilde{g}_1^* - g_1^*)} \cdot \frac{\tilde{g}_2 - (\tilde{g}_2 - g_2)}{\sqrt{2}} \right\} \Gamma^{\pi*}(0) \lambda_1^\pi(0) \right] (1 + M/q_P^0),
\end{aligned} \tag{2.25}$$

and

$$\begin{aligned}
& \tau^{-1} \left\{ \begin{array}{l} A' \rightarrow P + \pi^- \\ A' \rightarrow N + \pi^0 \end{array} \right\} \\
& = \frac{16}{\pi} q_P^0 q_\pi^0 q m_\Lambda^{-1} \\
& \times \left[\left| \left\{ \frac{\bar{g}_5'}{\tilde{g}_5' - (\tilde{g}_5' - g_5')} \right\} \Gamma^\pi(0) \right|^2 + \left| \left\{ \frac{\bar{g}_4'}{\tilde{g}_4' - (\tilde{g}_4' - g_4')} \right\} \lambda_1^\pi(0) \right|^2 \right. \\
& \left. + 2\Re i \left\{ \frac{\bar{g}_5'^* \tilde{g}_4'}{\tilde{g}_5'^* - (\tilde{g}_5'^* - g_5'^*)} \cdot \frac{\tilde{g}_4' - (\tilde{g}_4' - g_4')}{\sqrt{2}} \right\} \Gamma^{\pi*}(0) \lambda_1^\pi(0) \right] (1 - M/q_P^0).
\end{aligned} \tag{2.26}$$

Inserting the observed value 3×10^{-10} sec for both $\tau_{\Lambda \rightarrow P + \pi^-}$ and $\tau_{\Lambda' \rightarrow P + \pi^-}$ into (2.25) and (2.26), we obtain the following results:

a) In the case of $\bar{g}_1 = 0$ and $\bar{g}_2 = g$,

$$|\lambda_1^\pi(0)| = 0.69 \times 10^{-2} M^{3/2}. \tag{2.27}$$

b) In the case of $\bar{g}_2 = 0$ and $\bar{g}_1 = g$,

$$|\Gamma^\pi(0)| = 0.69 \times 10^{-2} M^{3/2}. \tag{2.28}$$

a)' In the case of $\bar{g}_5' = 0$ and $\bar{g}_4' = g$,

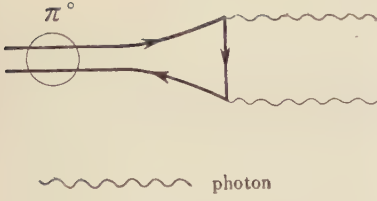
$$|\lambda_1^\pi(0)| = 0.98 \times 10^{-1} M^{3/2}. \tag{2.29}$$

b)' In the case of $\bar{g}_4' = 0$ and $\bar{g}_5' = g$,

$$|\Gamma^\pi(0)| = 0.98 \times 10^{-1} M^{3/2}. \tag{2.30}$$

vi) $\pi^0 \rightarrow 2\gamma$

Finally, let us consider the disintegration of π^0 -meson into two photons. This is considered as the decay of $P\bar{P}$ component of π^0 -meson ($(1/\sqrt{2})(N\bar{N} - P\bar{P}; {}^1S_0\text{-state})$) through the interaction,

Fig. 6. $\pi^0 \rightarrow 2\gamma$

$$H' = ie \bar{P} \gamma_\mu P A_\mu, \quad (2.31)$$

where A_μ 's denote a photon field.

This process is illustrated by the diagram in Fig. 6.

The matrix element of the lowest order for this process is given by

$$\bar{H}_{fi} = \langle 2\gamma | \bar{H} \frac{1}{H - m_\pi} \bar{H} | \pi^0 \rangle, \quad (2.32)$$

where \bar{H} consists of the free Hamiltonian of nucleon and photon fields and the strong interaction Hamiltonian (2.1) in I. Using eigenstate $|n\rangle$ of \bar{H} and its eigenvalue E_n , we can rewrite \bar{H}_{fi} in the form,

$$\bar{H}_{fi} = \sum_n \langle 2\gamma | \bar{H} | n \rangle \langle n | \bar{H} | \pi^0 \rangle / (E_n - m_\pi). \quad (2.33)$$

In order to facilitate the actual calculation, we shall make in \bar{H}_{fi} the approximation where \bar{H} is replaced by the free Hamiltonian of nucleon and photon field, and thus $|n\rangle$ and E_n are also replaced by the eigenstate of this free Hamiltonian and its eigenvalue, respectively. Thus, \bar{H}_{fi} becomes

$$\begin{aligned} \bar{H}_{fi} &= 2\sqrt{2} i e^2 / m_\pi \\ &\times \frac{1}{(2\pi)^3} \int d^3r d^3l \left[E(l) + E(l-q) - \frac{m_\pi}{2} \right]^{-1} E^{-1}(l-q) e^{-iql \cdot r} \\ &\times \left[M H^\pi(\mathbf{r}) + i \left(\mathbf{q} - \frac{1}{i} \frac{\partial}{\partial \mathbf{r}} \right) \lambda_4^\pi(\mathbf{r}) \right] \cdot \mathbf{e}(1) \times \mathbf{e}(2), \end{aligned} \quad (2.34)$$

where

$$E^2(\mathbf{p}) = \mathbf{p}^2 + M^2,$$

\mathbf{q} : the emitted photon momentum,

$\mathbf{e}(1), \mathbf{e}(2)$: the spin polarization vectors of the two photons.

Using the approximation in (2.34) as

$$E(l) + E(l-q) - m_\pi/2 \simeq 2E(l-q),$$

and taking up only to the second term in the power series expansion of $e^{-iql \cdot r}$, we obtain the final result for the matrix element:

$$\begin{aligned} \bar{H}_{fi} &\simeq -2\sqrt{2} \frac{e^2}{m_\pi} \int d^3r U(\mathbf{r}) e^{-iql \cdot r} \\ &\times \left\{ \mathbf{q} \lambda_4^\pi(\mathbf{r}) - i(M H^\pi(\mathbf{r}) - \frac{\partial}{\partial \mathbf{r}} \lambda_4^\pi(\mathbf{r})) \mathbf{r}/r \right\} \cdot \mathbf{e}(1) \times \mathbf{e}(2) \\ &\simeq -2\sqrt{2} e^2 m_\pi^{-1} W \mathbf{q} \cdot \mathbf{e}(1) \times \mathbf{e}(2), \end{aligned} \quad (2.35)$$

where

$$U(r) = e^{-Mr}/4\pi r, \quad H^\pi(r) = \mathbf{H}^\pi(\mathbf{r}) \cdot \mathbf{r}/r,$$

and

$$\begin{aligned} W &= 4\pi \left[\int_0^\infty r^2 dr U(r) \lambda_4^\pi(r) - (1/3) \int_0^\infty r^3 dr U(r) (MH^\pi(r) - (\partial/\partial r) \lambda_4^\pi(r)) \right] \\ &= -(4\pi/3) \left[\int_0^\infty r^3 \lambda_4^\pi(r) (\partial/\partial r) U(r) dr + M \int_0^\infty r^3 H^\pi(r) U(r) dr \right]. \end{aligned} \tag{2.36}$$

Thus, the lifetime is given by

$$\tau^{-1} = (e^4/6\pi) m_\pi^2 W^2. \tag{2.37}$$

In order to make the rough estimate of the numerical relation between the lifetime and the value of the wave function at the origin, we shall take into account only the first term of W . Further, assuming a rapid damping form for $\lambda_4^\pi(r)$ in the region $r \approx 1/M$, we shall limit the integral region of (2.36) within a nucleon Compton wave length $1/M$ and replace $\lambda_4^\pi(r)$ in the integrand by the value at the origin $\lambda_4^\pi(0)$. Thus,

$$\tau^{-1} = \frac{e^4}{6\pi} m_\pi^2 |\lambda_4^\pi(0)|^2 \left| \frac{4\pi}{3} \int_0^{1/M} r^3 \frac{\partial}{\partial r} U(r) dr \right|^2. \tag{2.38}$$

Inserting the observed value 5×10^{-15} sec for τ into the above equation, we obtain

$$|\lambda_4^\pi(0)| = 2.7 \times 10^{-2} M^{3/2}. \tag{2.39}$$

Now, the main results from i), ii),and vi) are summarized in Table I.

Table I*)

| decaying particle | $ \mathfrak{F} $ | $ \mathfrak{G} $ | $ \lambda_4^\pi(0) $ $M^{3/2}$ | $ \Gamma^\pi(0) $ $M^{3/2}$ | $ \lambda_4^0(0) $ $M^{3/2}$ | $ \Gamma^0(0) $ $M^{3/2}$ |
|-------------------|------------------|------------------|-----------------------------------|--------------------------------|---------------------------------|------------------------------|
| $K_{\mu 2}$ | 1.1 | 1.1 | | | 4.7×10^{-3} | 1.0×10^{-3} |
| $K_{\mu 3}$ | | | | | | |
| θ^0 | | | $\rightarrow *1.0 \times 10^{-2}$ | $*1.0 \times 10^{-2}$ | | |
| π^+ | | | 0.75×10^{-2} | 0.60×10^{-2} | | |
| π^0 | | | 2.7×10^{-2} | | | |
| A | | | 0.69×10^{-2} | 0.69×10^{-2} | | |
| A' | | | 0.98×10^{-1} | 0.98×10^{-1} | | |

§ 3. Discussions

Considering the results derived from the standpoint of the universal coupling strength of weak Fermi-interactions summarized in Table I, we shall remark several points.

*) The values given for the wave functions in this Table are not so general, because we assumed that a suitable one of five coupling constants in the Fermi-interaction is equal to the universal weak Fermi-coupling constant g , but the others are zero.
Each value with an asterisk is calculated with the help of the value of \mathfrak{F} and \mathfrak{G} as shown by each arrow.

The value of \mathfrak{F} derived from the $K_{\mu 0}$ -event is nearly equal to its maximum value 1 and shows a favourable overlapping of both wave functions of π - and K -mesons. On the contrary, the values of wave functions of π - and K -mesons at the origin are considerably different. The way to amend this point is either to make the value of λ_1^0 and Γ^0 larger or the value of \mathfrak{F} smaller. However, the latter process gives rise to an unwanted effect on the value of $|\lambda_1^\pi(0)|$ derived from $\theta^0 \rightarrow \pi^+ + \pi^-$, because, as shown in Table I, the value of \mathfrak{F} is used to determine, in the process $\theta^0 \rightarrow \pi^+ + \pi^-$, the value of $|\lambda_1^\pi(0)|$, which is in fairly good agreement with the values in the other processes. Therefore, taking into consideration this situation and (2.10), we find that it is reasonable to assume $|g_2|$, $|g_4'| < g$ and $|g_1|$, $|g_5'| < g$ for H_N^Λ and $H_N^{\Lambda'}$.

To explain the long lifetime of K^+ meson compared with θ^0 -meson ($\tau_{K^+} \sim 10^{-8}$ sec, $\tau_{\theta^0} \sim 10^{-10}$ sec), two kinds of standpoints may be conceivable, i.e., one is to explain the decay of K^+ in a similar way to that of θ^0 through the interaction H_N^Λ as described in § 2, and the other is to assume that all the decay processes are subject to the iso-spin selection rule $|JI| = (1/2)^{s,0}$ and thus $K^+ \rightarrow \pi^+ + \pi^0$ belongs to the first forbidden process. If we adopt the former standpoint, the coupling constants of H_N^Λ should satisfy the relation (2.24). The selection rule $|JI| = 1/2$ corresponds to $g_i = \bar{g}_i$ in our model, as readily seen from the expression of H_N^Λ and (2.19). On the other hand, in the Goldhaber's model¹⁰⁾ we have to allow the interaction $g\phi_{0^0}\phi_\pi\phi_\pi$ and forbid the interaction $g\phi_{K^+}\phi_\pi\phi_\pi$ to hold the selection rule $|JI| = 1/2$. Such a treatment, however, seems to us very asymmetric.

As to the values of $|\lambda_1^\pi(0)|$ and $|\Gamma^\pi(0)|$, they are almost consistent for all the phenomena except $\Lambda' \rightarrow P + \pi^-$. The result with respect to the process $\pi^0 \rightarrow 2\gamma$ also seems nearly consistent with the others, as shown in Table I. However, in this case the agreement may not be so reliable because of some sensitive approximation used in the calculation of the lifetime of this process. With regard to $\Lambda' \rightarrow P + \pi^-$, if we use the plausible value $\sim 10^{-2} M^{3/2}$ for $|\lambda_1^\pi(0)|$ in the expression of its lifetime, (2.26), we obtain the lifetime of the order of 10^{-8} sec. The cause of the difference between the lifetimes of both Λ and Λ' is also easily understood from the forms of the effective interactions $g\bar{\Lambda}P\pi$ and $g\bar{\Lambda}'\gamma_5 P\pi$.

At the end of this paper we would like to add some remarks on the existence of two kinds of K -mesons. In the present paper we have discussed the decay of $K_0(N, \tilde{\Lambda}')$ -meson as well as that of $K_c(N, \tilde{\Lambda})$ -meson, taking into consideration the standpoint of the preceding paper, where the existence of two kinds of Λ -particles was assumed. However, as far as we are concerned the decays of new particles discussed in the present paper, there does not appear any essential difference between the roles of these two K -mesons. It is, of course, evident that the conclusions derived in this paper hold good, irrespectively of the assumption of the existence of two kinds of Λ -particle, with an exception of the conclusion for K_0 -meson.

One of the important processes which we have not discussed in the present paper may be the event $\tau \rightarrow 3\pi$. For this process Λ' -particle must play an important role, be-

cause the decay mode is allowed only for $K_0(N, \tilde{l}')$ -meson, and not for $K_e(N, \tilde{l})$ -meson. Since the disintegrations of l' and $K_0(N, \tilde{l}')$ -meson are caused by the same interaction H_N^V , it seems very interesting to investigate the correlation between the lifetimes of both particles, l' and $K_0(N, \tilde{l}')$ -meson. This problem will be discussed in the forthcoming paper with other processes, e.g., $\pi \rightarrow e + \nu + \gamma$, etc.

References

- 1) S. Tanaka, Prog. Theor. Phys. **16** (1956), 625.
- 2) G. Wentzel, Phys. Rev. **79** (1950), 710.
- 3) E. Fermi and C. N. Yang, Phys. Rev. **76** (1949), 1739.
- 4) H. M. Moseley and N. Rosen, Phys. Rev. **80** (1950), 177.
- 5) R. T. Finkelstein, Phys. Rev. **88** (1952), 555.
- 6) M. Yamada, Prog. Theor. Phys. **10** (1953), 252.
H. Takebe, S. Nakamura and M. Taketani, Prog. Theor. Phys. **14** (1955), 317.
- 7) M. Fierz, Zs. f. Phys. **104** (1937), 553.
- 8) G. Wentzel, Phys. Rev. **101** (1956), 1215.
- 9) G. Takeda, Phys. Rev. **101** (1956), 1547.
- 10) M. Goldhaber, Phys. Rev. **101**(1956), 433.

Nuclear Magnetic Relaxation in Antiferromagnetics, II

Tôru MORIYA

Department of Physics, Tokyo Metropolitan University, Tokyo

(Received August 22, 1956)

Theory of nuclear magnetic relaxation in antiferromagnetics treated in the preceding paper for non-magnetic ions is extended to the case of magnetic ions. Strong hyperfine interactions and quadrupole interactions are the distinctive feature of the problem. The line width and the relaxation rate are calculated by using the spin wave approximation at low temperatures and the model of Gaussian random modulation at high temperatures.

Order estimations of T_1 and T_2 are made for several substances, which predict that the nuclear resonance will be difficult to detect above the Curie point because of the broadening due to the hyperfine interaction, while the resonance will become detectable at sufficiently low temperatures where the low frequency components of the local field spectra decrease remarkably.

The effects of the quadrupole interaction are also discussed.

§ 1. Introduction

In the preceding paper¹⁾, referred to below as I, the writer has presented a theoretical treatment of the nuclear magnetic relaxation of non-magnetic atoms in antiferromagnetic substances. The predominant relaxation mechanism in that case is provided by magnetic dipolar field coming from electron spins modulated by exchange interaction. The relaxation time of protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ calculated by using the spin wave approximation agreed well with the experiment both in the order of magnitude and in the nature of temperature dependence. On the other hand, for fluorine resonance in MnF_2 , the same mechanism has proved to be insufficient for explaining the reported absence of the resonance, and this compelled the present author to expect a hyperfine interaction between F nuclei and the unbalanced valence electron spins on F ions. This anticipated unbalance of spins was considered to result from the existence of a partial covalency in the crystal and the asymmetric arrangement of Mn^{++} ions around each F^- ion in MnF_2 .

In the present article, we shall extend our consideration to the nuclei which belong to the magnetic ions. In this case a nucleus is subjected to a very strong magnetic field due to the hyperfine interaction which is 10^2 – 10^3 times as large as the dipolar field from different magnetic ions, and its order of magnitude is 10^5 – 10^6 gauss. We expect, therefore, high resonance frequencies, broad line widths, and short relaxation times, and in some cases even the impossibility of observing the resonance. However, there is an effect of exchange narrowing which favours the observation of the resonance. Though the fluctuating local magnetic field is very large, its effect on nuclei is much diminished by a very rapid modulation due to the exchange interaction. Among the frequency spectra

of the local magnetic field, only the low frequency components can contribute to the line width and the relaxation. As was seen in I, the low frequency components of the local field spectra decrease very rapidly at low temperatures as temperature is lowered, so that the observation of the resonance seems to be possible at least at very low temperatures.

The present circumstances of the exchange narrowing correspond to the case of extreme narrowing where several authors^{(2), (3)} pointed out for certain cases that the relation $T_1 = T_2$ holds. In our case this relation will cease to hold below the Curie point because of the anisotropic nature of the electron spin fluctuations. For protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, this relation should approximately hold though this contribution to the line width will be overwhelmed by the dipolar broadening due to the nuclear moments. On the other hand, for those nuclei which belong to the magnetic ions, we expect that there is a certain temperature range where not only the relaxation rate but also the line width originates predominantly from the hyperfine interaction.

In the following sections we shall at first give the formulas of T_1 and T_2 for the nuclei which belong to the magnetic ions. In treating the exchange interaction we shall make use of the spin wave approximation below the Curie point and the model of Gaussian random process above the Curie point, as we did in I. At first in § 2 the quadrupole effect is neglected for brevity. The order estimations given in § 3 predicts that though the observation of the resonance absorption will be difficult above the Curie point, it will become possible at very low temperatures. In § 4 we shall deal with the effect of the quadrupole splitting and show that this causes no serious change in the line width. Finally, some discussions about the results obtained and some remarks on future experiments will be given in § 5.

§ 2. Formulas for line width and relaxation time

We shall here deal with a simple case where quadrupole interaction is absent. The magnetic field at a nucleus is produced not only by electrons on the nucleus but also by electronic spins on the other ions. The former contribution (hyperfine interaction) is much larger than the latter one (dipolar interaction). Hence, we may safely neglect the latter contribution, whose order of magnitude may, however, be inferred from the results of I.

We assume with considerable generality that the principal axes of the anisotropy of the electron spins and those of the hyperfine interaction coincide, and denote them as x , y , z . In the antiferromagnetic state the electronic spins are assumed to point in the $\pm z$ directions. The Hamiltonian of our problem is

$$\begin{aligned}\mathfrak{H} &= \mathfrak{H}_0 + \mathfrak{H}_e + \mathfrak{H}', \\ \mathfrak{H}_0 &= -\gamma_N \hbar \mathbf{I} \cdot \mathbf{H} + A_1 I_x \langle S_x \rangle + A_2 I_y \langle S_y \rangle + A_3 I_z \langle S_z \rangle, \\ \mathfrak{H}' &= A_1 I_x \delta S_x + A_2 I_y \delta S_y + A_3 I_z \delta S_z,\end{aligned}\tag{2.1}$$

with

$$\delta \mathbf{S} = \mathbf{S} - \langle \mathbf{S} \rangle,$$

where \mathfrak{H}_e denotes the Hamiltonian of the electron spin system which includes the exchange, dipole-dipole and anisotropic magneto-crystalline interactions, \mathbf{H} the external field, A_1 , A_2 and A_3 the principal values along the x , y and z directions of the hyperfine interaction tensor, $\langle \mathbf{S} \rangle$ its thermal average, \mathbf{I} the nuclear spin, and γ_N the gyromagnetic ratio of the nucleus.

\mathfrak{H}_0 determines the Zeeman levels of a nucleus, \mathfrak{H}_e modulates \mathfrak{H}' , and \mathfrak{H}' thus modulated contributes to the line width and the relaxation rate. We shall at first diagonalize \mathfrak{H}_0 by the following orthogonal transformation of the coordinates system :

$$\begin{aligned} x &= \alpha_1 \xi + \alpha_2 \eta + \alpha_3 \zeta, \\ y &= \beta_1 \xi + \beta_2 \eta + \beta_3 \zeta, \\ z &= \gamma_1 \xi + \gamma_2 \eta + \gamma_3 \zeta, \end{aligned} \quad (2.2)$$

where α , β , γ satisfy

$$\alpha : \beta : \gamma = (\gamma_N \hbar H_x - A_1 \langle S_x \rangle) : (\gamma_N \hbar H_y - A_2 \langle S_y \rangle) : (\gamma_N \hbar H_z - A_3 \langle S_z \rangle). \quad (2.3)$$

Hamiltonian (2.1) then becomes as follows :

$$\mathfrak{H}_0 = -\hbar \omega_0 I_z$$

with

$$\hbar \omega_0 = [(\gamma_N \hbar H_x - A_1 \langle S_x \rangle)^2 + (\gamma_N \hbar H_y - A_2 \langle S_y \rangle)^2 + (\gamma_N \hbar H_z - A_3 \langle S_z \rangle)^2]^{1/2}, \quad (2.4)$$

$$\mathfrak{H}' = I_z K_0 + \frac{1}{2} (I_+ K_- + I_- K_+)$$

with

$$K_0 = \frac{\alpha A_1 - i\beta A_2}{2} \delta S_+ + \frac{\alpha A_1 + i\beta A_2}{2} \delta S_- + \gamma A_3 \delta S_z, \quad (2.5)$$

$$K_{\pm} = \frac{\alpha_{\pm} A_1 - i\beta_{\pm} A_2}{2} \delta S_+ + \frac{\alpha_{\pm} A_1 + i\beta_{\pm} A_2}{2} \delta S_- + \gamma_{\pm} A_3 \delta S_z,$$

$$I_{\pm} = I_x \pm iI_y, \quad S_{\pm} = S_x \pm iS_y, \quad (2.6)$$

$$\alpha_{\pm} = \alpha_1 \pm i\alpha_2, \quad \beta_{\pm} = \beta_1 \pm i\beta_2, \quad \gamma_{\pm} = \gamma_1 \pm i\gamma_2.$$

Now, we shall calculate the line width and the relaxation time. This is a problem of exchange narrowing. A theoretical treatment of exchange narrowing was given by Anderson and Weiss¹⁾ as well as by Kubo and Tomita.³⁾ The former deals with the line shape and is based on the adiabatic approximation and the model of Gaussian random process, while the latter is a perturbational approach including explicitly the adiabatic and non-adiabatic effects as well. A general formula for T_1 is also given by the latter. We shall here use chiefly the latter method.

The line contour is obtained as the Fourier transform of the correlation function, which can be calculated by perturbation method following Kubo-Tomita's procedure. We

get for the line contour

$$I(\omega) = \int_{-\infty}^{\infty} e^{i\omega t - \psi(t)} dt, \quad (2.7)$$

$$\psi(t) = \sigma_0^2 \int_0^t d\tau (t-\tau) f_0(\tau) + \sigma_1^2 \int_0^t d\tau (t-\tau) e^{i\omega_0 \tau} f_1(\tau), \quad (2.8)$$

with

$$\begin{aligned} \sigma_0^2 &= \langle K_0^2 \rangle / \hbar^2, \\ f_0(t) &= \langle \{K_0(t) K_0\} \rangle / \langle K_0^2 \rangle, \\ &= [\frac{1}{2} (\alpha^2 A_1^2 + \beta^2 A_2^2) \langle \{\partial S_+(t) \partial S_-\} \rangle + \gamma^2 A_3^2 \langle \{\partial S_z(t) \partial S_z\} \rangle] / \langle K_0^2 \rangle, \\ \sigma_1^2 &= \langle \{K_+ K_-\} \rangle / 2\hbar^2, \\ f_1(t) &= \langle \{K_+(t) K_-\} \rangle / \langle \{K_+ K_-\} \rangle \\ &= [\frac{1}{2} \{ (1-\alpha^2) A_1^2 + (1-\beta^2) A_2^2 \} \langle \{\partial S_+(t) \partial S_-\} \rangle \\ &\quad + (1-\gamma^2) A_3^2 \langle \{\partial S_z(t) \partial S_z\} \rangle] / \langle \{K_+ K_-\} \rangle, \end{aligned} \quad (2.9)$$

where

$$K(t) = e^{it\mathfrak{H}_0/\hbar} K e^{-it\mathfrak{H}_0/\hbar}.$$

Here $\langle \rangle$ means the statistical average, $\{AB\}$ the symmetrized product of A and B . When $f_0(t)$ and $f_1(t)$ decay rapidly as t increases, i.e. their characteristic decay time τ_c satisfies $\omega_0 \tau_c \ll 1$, we may approximate for the line centre part as follows:

$$\psi(t) = (\sigma_0^2 \tau_0 + \sigma_1^2 \tau_1) |t|, \quad (2.10)$$

$$\begin{aligned} \tau_0 &= \int_0^{\infty} f_0(\tau) d\tau, \\ \tau_1 &= \int_0^{\infty} f_1(\tau) d\tau. \end{aligned} \quad (2.11)$$

This gives a Lorentzian line shape whose half width is

$$1/T_2 = (\mathcal{A}\omega)_{1/2} = 1/T_2' + 1/T_1', \quad (2.12)$$

where

$$1/T_2' = \sigma_0^2 \tau_0, \quad 1/T_1' = \sigma_1^2 \tau_1. \quad (2.13)$$

$1/T_2'$ is the line width due to the secular part of the perturbing Hamiltonian and $1/T_1'$ is that due to the non-secular part.

The thermal relaxation time T_1 is given by

$$1/T_1 = \sigma_1^2 \int_{-\infty}^{\infty} dt \cos \omega_0 t f_1(t). \quad (2.14)$$

When the narrowing is extreme, i.e. $\omega_0\tau_1 \ll 1$, this can be simplified as

$$1/T_1 = 2\sigma_1^2\tau_1 = 2/T_1'. \quad (2.15)$$

This means that $1/T_1$ is just twice the non-secular line width $1/T_1'$ in the case of extreme narrowing. The formulas (2.7) — (2.15) will be applied below to individual cases.

(1) *paramagnetic region* $T > T_c$

In the paramagnetic region, where $\partial S \simeq S$, the model of Gaussian random modulation seems to be reasonable.⁽⁴⁾ Basing on this model we calculated in I the correlation functions of the electronic spin of an ion at high temperatures. The result at sufficiently high temperatures was

$$\begin{aligned} \langle S_z(t) S_z \rangle &= \frac{1}{2} \langle S_+(t) S_- \rangle = \frac{1}{2} \langle S_-(t) S_+ \rangle \\ &= (1/3) S(S+1) \exp(-\frac{1}{2} \omega_e^2 t^2) \end{aligned} \quad (2.16)$$

with

$$\omega_e^2 = 2J^2 S(S+1) / 3\hbar^2, \quad (2.17)$$

where J is the exchange integral times -2 , and z the number of nearest neighboring magnetic ions. Substituting (2.16) in (2.9) we get

$$\begin{aligned} \sigma_0^2 &= S(S+1) (\alpha^2 A_1^2 + \beta^2 A_2^2 + \gamma^2 A_3^2) / 3\hbar^2, \\ \sigma_1^2 &= S(S+1) \{ (1-\alpha^2) A_1^2 + (1-\beta^2) A_2^2 + (1-\gamma^2) A_3^2 \} / 6\hbar^2, \\ f_0(\tau) &= f_1(\tau) = \exp(-\frac{1}{2} \omega_e^2 \tau^2). \end{aligned} \quad (2.18)$$

As the condition $\omega_e \gg \omega_0$ is usually satisfied, we can use (2.12), (2.13) and (2.15) for the calculation of T_1 and T_2 . The results are

$$\frac{1}{T_2} = \frac{\sqrt{\pi}/2S(S+1)}{6\hbar^2\omega_e} \{ (1+\alpha^2) A_1^2 + (1+\beta^2) A_2^2 + (1+\gamma^2) A_3^2 \}, \quad (2.19)$$

$$\frac{1}{T_1} = \frac{\sqrt{\pi}/2S(S+1)}{3\hbar^2\omega_e} \{ (1-\alpha^2) A_1^2 + (1-\beta^2) A_2^2 + (1-\gamma^2) A_3^2 \}. \quad (2.20)$$

When the hyperfine interaction is isotropic, we get

$$T_1 = T_2. \quad (2.21)$$

This holds also when we take directional averages of (2.19) and (2.20).

(2) *antiferromagnetic region* $T < T_c$

Below the Curie temperature an antiferromagnetic ordered arrangement of spins takes place. At low temperatures, spin deviations can well be described by the spin wave approximation and the correlation functions, $\langle \{ \partial S_z(t) \partial S_z \} \rangle$ and $\langle \{ \partial S_+(t) \partial S_- \} \rangle$, are obtained in a closed form or in the form of the Fourier transform of the frequency spectra which can explicitly be written. As was already seen in I, the frequency spectra

can be divided into two parts, namely, the low frequency part and the high frequency part; the former arises from the interference of two spin waves, while the latter consists mainly of the spin wave spectra. Both of these two components are functions of temperature; the former vanishes at 0 K while the latter does not, since the zero point motion contributes to the latter.

The high frequency components are all located in the region of higher frequencies than the lowest spin wave frequency $\omega_{\min} = 2S\sqrt{z}JK/\hbar$ (K : anisotropy constant), which is far larger than $\gamma_N A |\partial S|$, so that they make no contributions to the line width.

The low frequency components make a heap widely spread around zero frequency. Therefore, the Fourier transform of the low frequency part decays rapidly with the lapse of time, so that the argument for the case of extreme narrowing may be applicable. We get the following results (See Appendix):

$$\phi(t) = (1/T_2)|t| = (1/T_2' + 1/T_1')|t|, \quad (2.22)$$

$$1/T_2' = S(S+1)[(\alpha^2 A_1^2 + \beta^2 A_2^2)\tau_1' + \gamma^2 A_3^2 \tau_0']/3\hbar^2, \quad (2.23)$$

$$1/T_1' = 1/2T_1 = S(S+1)[\{(1-\alpha^2)A_1^2 + (1-\beta^2)A_2^2\}\tau_1' + (1-\gamma^2)A_3^2 \tau_0']/6\hbar^2, \quad (2.24)$$

where

$$\begin{aligned} \tau_0' &= \{3/S(S+1)\} \int_0^\infty \langle \{\partial S_z(t) \partial S_z\} \rangle dt, \\ \tau_1' &= \{3/2S(S+1)\} \int_0^\infty \langle \{\partial S_+(t) \partial S_-\} \rangle dt. \end{aligned} \quad (2.25)$$

The relation $1/T_1 = 2/T_1'$ remains valid also below the Curie point.

Now, we have only to calculate τ_0' and τ_1' . The correlation functions of the spin deviation were already calculated in I with the use of the spin wave approximation. If we take account of terms up to the fourth order with respect to the spin wave creation and annihilation operators, $\langle \{\partial S_z(t) \partial S_z\} \rangle$ has low frequency components while $\langle \{\partial S_+(t) \partial S_-\} \rangle$ has not; the low frequency components of the latter appear in the sixth order terms. We shall neglect the latter. Then τ_1' vanishes and the following results are obtained:

$$1/T_2 = A_3^2 S(S+1) \tau_0' (1+\gamma^2)/6\hbar^2, \quad (2.26)$$

$$1/T_1 = A_3^2 S(S+1) \tau_0' (1-\gamma^2)/3\hbar^2, \quad (2.27)$$

where, from the result of I,

$$\tau_0' = \frac{3\pi}{2S(S+1)} \int d\omega \{N(\omega)\}^2 \frac{e^{\beta\hbar\omega/kT}}{(e^{\beta\hbar\omega/kT} - 1)^2} \left[1 + \left\{ \frac{(1+\alpha)zJS}{\hbar\omega} \right\}^2 \right]. \quad (2.28)$$

Here $N(\omega)$ denotes the normalized state density of spin waves and α is defined by $\alpha = 2K/zJ$. The expression (2.28) was obtained by using the two-sublattice model with uniaxial anisotropy. In the previous paper, I, we have calculated (2.28) adopting the long wave approximation for the spin waves.

§ 3. Order estimation of T_2 and T_1

We shall now estimate the order of magnitude of T_2 and T_1 for some antiferromagnetic substances making use of the formulas derived in the previous section. We neglected in the previous section the effect of the quadrupole interaction, though actually any nucleus of iron group elements has either both magnetic dipole and electric quadrupole moments or none of them. Accordingly, it seems at first sight that the formulas derived in the previous section cannot be directly applied to the real cases. However, as we shall see in the next section, the order of magnitude of the line width $1/T_2$ is not appreciably changed by taking account of the quadrupole splitting, and $1/T_1$ does not exceed $1/T_2$.

The order estimation given below will be of use in predicting whether the observation of the resonance absorption is possible or not.

(1). paramagnetic region $T > T_c$

From the formulas (2.19) and (2.20), taking the average over directions, we get

$$1/T_1 = 1/T_2 = \sqrt{2\pi} (A_1^2 + A_2^2 + A_3^2) S(S+1) / 9\hbar^2 \omega_e. \quad (3.1)$$

The hyperfine coupling constant A may be estimated from hyperfine structure data in the paramagnetic resonance on magnetically diluted crystals⁵⁾. It is because the hyperfine coupling is chiefly of atomic nature and is not much affected by surroundings when placed in a crystal. Estimated values are as follows:

$$\begin{array}{lll} \text{Mn}^{++} & A = 0.009 \text{ cm}^{-1} & \text{isotropic,} \\ \text{Co}^{++} & A_{\text{max}} = 0.03 \text{ cm}^{-1} & \text{anisotropic,} \\ \text{Cu}^{++} & A_{\text{max}} = 0.01 \text{ cm}^{-1} & \text{anisotropic.} \end{array} \quad (3.2)$$

Then, from (2.29) we get

$$\begin{aligned} (1/T_2)_{\text{Mn}} &= 2 \times 10^{19} \omega_e^{-1} \text{ sec}^{-1}, \\ (1/T_2)_{\text{Co}} &= 5 \times 10^{19} \omega_e^{-1} \text{ sec}^{-1}, \\ (1/T_2)_{\text{Cu}} &= 1 \times 10^{18} \omega_e^{-1} \text{ sec}^{-1}. \end{aligned} \quad (3.3)$$

ω_e can be obtained from (2.17) estimating the exchange integral from the Curie point or the perpendicular susceptibility with the use of the molecular field theory. Estimated values of ω_e and $T_1 = T_2$ are tabulated in Table 1. We see that the expected order of magnitude of T_1 and T_2 is, for any substance, shorter than 10^{-6} sec. and accordingly it seems that the observation of the resonance absorption is impossible or at least very difficult in the paramagnetic region.

Table 1. Relaxation times of metal nuclei in several antiferromagnetic substances at high temperatures.

| Substance | ω_e | $1/T_2 = 1/T_1 (\text{sec}^{-1})$ |
|--------------------------------------|----------------------|-----------------------------------|
| MnO | 5×10^{12} | 4×10^6 |
| MnF ₂ | 2×10^{12} | 1×10^7 |
| CoO | 1.5×10^{13} | 3×10^6 |
| CoF ₂ | 2×10^{12} | 2.5×10^7 |
| CuCl ₂ ·2H ₂ O | 2×10^{12} | 5×10^5 |

(2) *antiferromagnetic region* $T < T_c$

We may use Eqs. (2.26) and (2.27). According to these formulas, T_2 does not strongly depend on the direction of the magnetic field at the nucleus, while T_1 depends markedly on it. Under usual conditions ($\gamma \simeq 1$) T_2 will be far shorter than T_1 . As will be stated in the next section, the relation $T_2 \lesssim T_1$ holds generally even when the quadrupole splitting is taken into account. Accordingly, whether the observation of the resonance is possible or not will depend on the magnitude of T_2 unless T_1 is so long as to make the resonance unobservable. We shall estimate T_2 below. Using the values of (2.30) and putting $\gamma=1$ in (2.26) we get

$$\begin{aligned} (1/T_2)_{\text{Mn}} &= 0.8 \times 10^{10} \tau_0', \\ (1/T_2)_{\text{Co}} &= 4 \times 10^{10} \tau_0', \\ (1/T_2)_{\text{Cu}} &= 0.9 \times 10^{18} \tau_0'. \end{aligned} \quad (3.4)$$

τ_0' is given by (2.28). In the preceding paper, I, we calculated τ_0' for $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ and MnF_2 as functions of temperature*. We can here make use of them. T_2 for Mn in MnF_2 calculated in this way is shown in Table 2a. For $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ we shall estimate $(T_2)_{\text{Cu}}$ in another way, namely, by correlating it to the relaxation time T_1 for protons whose experimental values are known.⁶⁾ According to the present notation, T_1 for protons is written as

$$(1/T_1)_{\text{proton}} = (1/4) \gamma_e^2 \gamma_N^2 \hbar^2 \sum_l r_l^{-6} F_l(\alpha \beta \gamma) \tau_0',$$

where γ_e denotes the gyromagnetic ratio of an electron, r_l the distance between a proton and the l -th Cu ion and $F_l(\alpha \beta \gamma)$ a certain function of the direction of the local magnetic field at the nucleus and of \mathbf{r}_l , the position vector of the l -th Cu ion relative to the nucleus, and is of order unity. Evaluating numerical values for these parameters we get

$$(1/T_1)_{\text{proton}} = 2.5 \times 10^4 \tau_0'.$$

Comparing this with (2.32), we obtain

$$(T_2)_{\text{Cu}} / (T_1)_{\text{proton}} \sim 3 \times 10^{-4}.$$

From this and the experimental values of T_1 for protons,⁶⁾ T_2 for Cu can be estimated. The result is shown in Table 2b.

Table 2a. Line width of a Mn-resonance line in MnF_2 below T_c

| $T^\circ\text{K}$ | τ_0' (sec) | $1/T_2$ (sec^{-1}) |
|-------------------|---------------------|-------------------------------|
| 1.5 | 7×10^{-22} | 6×10^{-5} |
| 7.5 | 7×10^{-17} | 6×10^2 |
| 15 | 2×10^{-16} | 2×10^3 |
| 30 | 2×10^{-15} | 2×10^4 |
| 45 | 8×10^{-15} | 7×10^4 |

Table 2b. Line width of a Cu-resonance line in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ below T_c .

| $T^\circ\text{K}$ | $1/T_2$ (sec^{-1}) |
|-------------------|-------------------------------|
| 1 | 2×10^2 |
| 1.5 | 5×10^3 |
| 2 | 3×10^4 |
| 2.5 | 1.5×10^5 |
| 3 | 3×10^5 |

* Strictly speaking, we calculated in I a certain equivalent of τ_0' .

These order estimations predict that at very low temperatures the line width is narrow enough to permit experimental observations. One should not here lay stress upon the figures in the tables; they are only of order of magnitude significance.

§ 4. Effect of quadrupole interaction

As we have already stated, any nucleus of iron group elements which has a magnetic moment has also an electric quadrupole moment. Accordingly the resonance spectrum splits usually into several lines. The electric field gradient to which the nucleus is subjected may be considered to be caused mainly by *d*-electrons. If we neglect the other origins, both the electric field gradient tensor ∇E and the hyperfine interaction tensor will have the common principal axes. This statement is also valid when the contribution of the crystalline field to ∇E is taken into account unless the crystalline field has unusually low symmetry such as in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$.

The calculation of the line width of each split line is rather complicated. However, the broadening due to the secular perturbations can easily be calculated while the non-secular effect is of the same nature as the thermal relaxation effect which can easily be treated. We may regard the latter as a measure of the non-secular broadening.

As for the thermal relaxation, a single characteristic time will in this case not be sufficient to describe the whole relaxation phenomena. We calculate here the thermal transition probabilities corresponding to the transitions between any two levels.

We shall now deal with the following two cases, where temperature is always lower than the Curie point:

(1) *The case where the principal axes of magnetic anisotropy, hyperfine interaction and quadrupole interaction coincide. (MnF_2 , CoF_2 , etc.)*

We assume that x , y , z axes are the common principal axes and the easy direction is along the z axis. The Hamiltonian is written as

$$\begin{aligned} \mathfrak{H}_0 = & (A_1 \langle S_x \rangle - \gamma_N \hbar H_x) I_x + (A_2 \langle S_y \rangle - \gamma_N \hbar H_y) I_y \\ & + (A_3 \langle S_z \rangle - \gamma_N \hbar H_z) I_z + P_x I_x^2 + P_y I_y^2 + P_z I_z^2, \end{aligned} \quad (4.1)$$

where P_x , P_y and P_z are the quadrupole coupling constants:

$$P_x = -(eQ/2I(2I-1)) \partial E_x / \partial x, \quad \text{etc.}$$

With the abbreviations:

$$\begin{aligned} A_3 \langle S_z \rangle - \gamma_N \hbar H_z &= B_0, \\ (A_1 \langle S_x \rangle - \gamma_N \hbar H_x) \pm i(A_2 \langle S_y \rangle - \gamma_N \hbar H_y) &= B_{\pm}, \\ P_z/2 &= P, \quad (P_x - P_y)/P_z = \eta, \end{aligned} \quad (4.2)$$

(4.1) becomes

$$\mathfrak{H}_0 = B_0 I_z + \frac{1}{2} (B_- I_+ + B_+ I_-) + P [3I_z^2 + \frac{1}{2} \eta (I_+^2 + I_-^2)]. \quad (4.3)$$

The perturbing Hamiltonian is

$$\mathfrak{H}' = A_0 \partial S_z I_z + \frac{1}{4} (A_- \partial S_+ + A_+ \partial S_-) I_+ + \frac{1}{4} (A_+ \partial S_+ + A_- \partial S_-) I_-, \quad (4.4)$$

where

$$A_{\pm} = A_1 \pm A_2.$$

At first we shall calculate the eigen-values and the eigen-functions of (4.3). As the electron spins point almost in the $\pm z$ directions, the first term will be greater than all the others. The quadrupole interaction is usually smaller than the hyperfine interaction by a factor of about 10. We may use the perturbation calculation regarding the first term as an unperturbed Hamiltonian and the other terms as perturbations. We get

$$E_m = B_0 m + 3 P m^2 + (B_+ B_- / 2 B_0) m + (\gamma^2 P^2 / B_0) [-m^3 + m \{I(I+1) - \frac{1}{2}\}], \quad (4.5)$$

$$\begin{aligned} \Psi_m = & \phi_m + (B_+ / 2 B_0) \sqrt{(I+m)(I-m+1)} \phi_{m-1} - (B_- / 2 B_0) \sqrt{(I-m)(I+m+1)} \phi_{m+1} \\ & + (\gamma P / 4 B_0) \sqrt{(I+m)(I+m+1)(I-m+1)(I-m+2)} \phi_{m-2} \\ & - (\gamma P / 4 B_0) \sqrt{(I-m)(I-m+1)(I+m+1)(I+m+2)} \phi_{m+2}, \end{aligned} \quad (4.6)$$

where ϕ_m means the nuclear spin function corresponding to the state $I_z = m$.

When a r.f. field is applied, transitions corresponding to $|Jm| = 1, 2, 3$ are generally all possible. The intensities of them depend on B_0 , B_{\pm} , p , γ and the direction of the polarization of the r.f. field. When $B_0 \cdot B_{\pm}$ and $B_0 \cdot P$ are satisfied, as is usually the case, the transitions corresponding to $|Jm| = 1$ overcome the others. We shall here consider only these transitions.

The line broadening due to the secular perturbations can be treated by the Anderson-Weiss' theory as well as by the Kubo-Tomita's theory. The line contour corresponding to the transition $m \rightarrow m'$ is given by

$$I_{mm'}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t - \phi_{mm'}(t)} dt, \quad (4.7)$$

$$\phi_{mm'}(t) = \frac{1}{\hbar^2} \int_0^t d\tau (t - \tau) \langle \{\mathfrak{H}'_{m'm'}(\tau) - \mathfrak{H}'_{mm}(\tau)\} \{\mathfrak{H}'_{m'm'} - \mathfrak{H}'_{mm}\} \rangle. \quad (4.8)$$

From (4.4) and (4.6) we calculate

$$\begin{aligned} \mathfrak{H}'_{mm} - \mathfrak{H}'_{m-1,m-1} = & A_0 \partial S_z [1 + (B_+ B_- / 2 B_0^2) F_1(I, m) + (\gamma^2 P^2 / 8 B_0^2) F_2(I, m)] \\ & + [(B_+ A_- + B_- A_+) \partial S_+ + (B_+ A_+ + B_- A_-) \partial S_-] / 4 B_0 \\ & + \gamma P [(B_+ A_+ + B_- A_-) \partial S_+ + (B_+ A_- + B_- A_+) \partial S_-] F_3(I, m) / 16 B_0^2, \end{aligned} \quad (4.9)$$

where

$$\begin{aligned} F_1(I, m) &= I(I+1) - 3m(m+1), \\ F_2(I, m) &= I^2(I+1)^2 - 12I(I+1) + 18 \\ &\quad - m(m-1)[44 - 6I(I+1) + 5m(m-1)], \\ F_3(I, m) &= 2(2m-1)[3 - I(I+1) + m(m-1)]. \end{aligned} \quad (4.10)$$

From (4.8) and (4.9), neglecting τ_1' as before, we get

$$\phi_{mm-1}(t) = |t|/T_2', \quad (4.11)$$

$$1/T_2' = \{A_3^2 S(S+1) \tau_0' / 3\hbar^2\} [1 + (B_+ B_- / B_0^2) F_1(I, m) + (\eta^2 P^2 / 4B_0^2) F_2(I, m)], \quad (4.12)$$

where $1/T_2'$ is the line width due to the secular effect and τ_0' is defined by (2.25). Comparing this result with § 2, we see that the effect of the external field and the quadrupole interaction on the line width appears as a multiplicative factor

$$1 + (B_+ B_- / B_0^2) F_1(I, m) + (\eta^2 P^2 / 4B_0^2) F_2(I, m).$$

Therefore, when the quadrupole interaction is far smaller than the hyperfine interaction the line width of each split line is not remarkably changed from that before the splitting, though the relative line width changes from line to line since $F_1(I, m)$ and $F_2(I, m)$ take positive as well as negative values. Even when the quadrupole interaction is not small and the perturbation calculation is not justified, the multiplicative correction factor will remain of order unity.

Next we shall turn to the thermal transition probability. The following formula is useful:

$$W_{m \rightarrow m'} = -\frac{1}{\hbar^2} \int_{-\infty}^{\infty} dt \langle \mathfrak{S}'_{mm'}(t) \mathfrak{S}'_{m'm}(0) \rangle \quad (4.13)$$

where

$$\mathfrak{S}'(t) = e^{it(\mathfrak{S}_0 + \mathfrak{S}_c)/\hbar} \mathfrak{S}' e^{-it(\mathfrak{S}_0 + \mathfrak{S}_c)/\hbar}.$$

From (4.4), (4.5) and (4.6) we calculate the transition probability corresponding to $m \rightarrow m-1$. The result is

$$W_{m \rightarrow m-1} = [A_3^2 S(S+1) \tau_0' (I+m)(I-m+1) / 6\hbar^2] (B_+ B_- / B_0^2) [1 - \{(B_+^2 + B_-^2) \eta P / 4B_+ B_- B_0\} (2m-1) \{I(I+1) - m(m-1) - 4\}]. \quad (4.14)$$

Here τ_1' was again neglected. We see from this, referring to § 2, that the effect of the quadrupole interaction appears as a multiplicative factor

$$1 - \{(B_+^2 + B_-^2) \eta P / 4B_+ B_- B_0\} (2m-1) \{I(I+1) - m(m-1) - 4\}.$$

This factor remains of order unity even if the quadrupole interaction is so large that the perturbation calculation is not justified.

When the external field is absent (4.14) vanishes. Even when there is an external field the factor $B_+ B_- / B_0^2$ is usually very small and accordingly (4.14) is very small. The actual thermal relaxation will be provided by other mechanisms, e.g. the dipolar field coming from the other ions or the low frequency part of $\langle \delta S_+(t) \delta S_- \rangle$ which was neglected above. Any way, we may expect in this case $T_1 \gg T_2$.

(2) The case of $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$

We shall here discuss on the Cu-resonance in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$. The crystal structure

has orthorhombic symmetry and the antiferromagnetic ordering occurs along the a -axis.⁷⁾ However, the crystalline field around a Cu^{++} ion has a lower symmetry and its principal axes do not coincide with the crystal axes.

We may take the common principal axes* for the hyperfine and quadrupole interactions and denote them as ξ , η , ζ . Considering the crystal symmetry, the b -axis proves to be a principal axis. The coordinate transformation between $x=b$, $y=c$, $z=a$ and ξ , η , ζ may be written as

$$\begin{aligned} x &= \xi, \\ y &= \eta \cos \alpha + \zeta \sin \alpha, \\ z &= -\eta \sin \alpha + \zeta \cos \alpha. \end{aligned} \quad (4.15)$$

These coordinate axes are illustrated in Fig. 1. There are two kinds of Cu^{++} sites in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ where the crystalline fields are the mirror images of each other with respect to bc -plane.

We neglect henceforth the external field so that the average spins point in the $\pm z$ directions. Hamiltonian is written as

$$\begin{aligned} \mathcal{H}_0 &= (A_\zeta I_\zeta \cos \alpha - A_\eta I_\eta \sin \alpha) \langle S_z \rangle \\ &+ P_\xi I_\xi^2 + P_\eta I_\eta^2 + P_\zeta I_\zeta^2, \end{aligned} \quad (4.16)$$

$$\begin{aligned} \mathcal{H}' &= (A_\zeta I_\zeta \cos \alpha - A_\eta I_\eta \sin \alpha) \delta S_z \\ &+ \frac{1}{2} \{ A_\xi I_\xi - i(A_\zeta I_\zeta \sin \alpha + A_\eta I_\eta \cos \alpha) \} \delta S_+ \\ &+ \frac{1}{2} \{ A_\xi I_\xi + i(A_\zeta I_\zeta \sin \alpha + A_\eta I_\eta \cos \alpha) \} \delta S_-. \end{aligned} \quad (4.17)$$

With another coordinate transformation:

$$\begin{aligned} I_\xi &= I_X, \\ I_\eta &= I_Y \cos \beta - I_Z \sin \beta, \\ I_\zeta &= I_Y \sin \beta + I_Z \cos \beta, \end{aligned} \quad (4.18)$$

with

$$\tan \beta = A_\eta \sin \alpha / A_\zeta \cos \alpha, \quad (4.19)$$

we get

$$\begin{aligned} \mathcal{H}_0 &= A_1 \langle S_z \rangle I_Z \\ &+ Q_0 I_Z^2 + Q_1 \left\{ \frac{1}{2} (I_Z I_+ + I_+ I_Z) - \frac{1}{2} (I_Z I_- + I_- I_Z) \right\} \\ &+ Q_2 \frac{1}{2} (I_+^2 + I_-^2), \end{aligned} \quad (4.20)$$

* This is correct when only the interaction of a nucleus with d -electrons is taken into account.

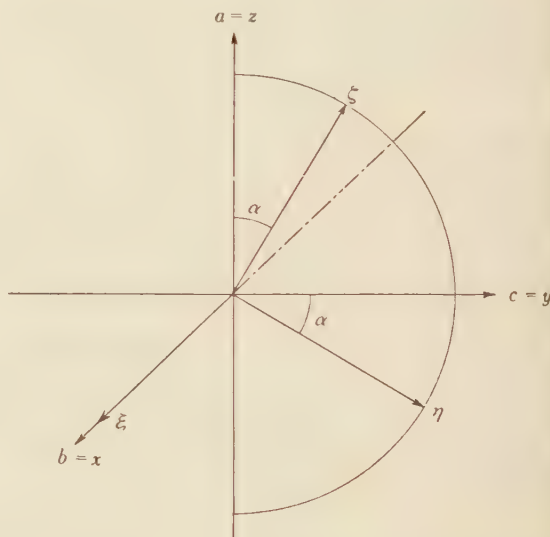


Fig. 1.

$$\begin{aligned} \mathfrak{H}' = & I_Z A_1 [\delta S_z - \frac{1}{2} i \lambda (\delta S_+ - \delta S_-)] \\ & + \frac{1}{2} I_+ (A_- \delta S_+ + A_+ \delta S_-) + \frac{1}{2} I_- (A_+ \delta S_+ + A_- \delta S_-), \end{aligned} \quad (4 \cdot 21)$$

where

$$\begin{aligned} A_1 &= (A_\eta^2 \sin^2 \alpha + A_\zeta^2 \cos^2 \alpha)^{1/2}, \\ A_\pm &= \frac{1}{2} (A_\mp \pm A_\eta A_\zeta / A_1), \\ Q_0 &= \frac{3}{2} (P_\eta \sin^2 \beta + P_\zeta \cos^2 \beta), \\ Q_1 &= i (P_\eta - P_\zeta) \sin \beta \cos \beta, \\ Q_2 &= P_\mp + \frac{1}{2} (P_\eta \sin^2 \beta + P_\zeta \cos^2 \beta), \\ \lambda &= (A_\zeta^2 - A_\eta^2) \sin \alpha \cos \alpha / A_1^2. \end{aligned} \quad (4 \cdot 22)$$

The eigenfunctions of \mathfrak{H}_0 are expressed in terms of the eigenfunctions of I_Z as follows ($I = \frac{3}{2}$ for Cu) :

$$\Psi_m = \sum_{m'} a_{mm'} \phi_{m'}, \quad m, m' = \pm 3/2, \pm 1/2, \quad (4 \cdot 23)$$

where

$$a_{3/2-3/2} = a_{1/2-1/2} = a_{-1/2 \ 1/2} = a_{-3/2 \ 3/2} = 0, \quad (4 \cdot 24)$$

$a_{mm'}$ and the eigenvalue E_m are obtained by solving the fourth order secular equation. Here we have recourse simply to the perturbation calculation. The results given below are up to the second order.

$$\begin{aligned} E_{\pm 3/2} &= \pm 3 \hbar \omega_0 / 2 + 9 Q_0 / 4 \mp 3 Q_1^2 / \hbar \omega_0 \pm 3 Q_2^2 / 2 \hbar \omega_0, \\ E_{\pm 1/2} &= \pm \frac{1}{2} \hbar \omega_0 + \frac{1}{4} Q_0 \pm 3 Q_1^2 / \hbar \omega_0 \pm 3 Q_2^2 / 2 \hbar \omega_0, \\ \hbar \omega_0 &= A_1 \langle S_z \rangle, \end{aligned} \quad (4 \cdot 25)$$

$$\begin{aligned} a_{mm} &= 1, \quad a_{1/2 \ 3/2} = a_{3/2 \ 1/2} = -a_{-1/2-3/2} = -a_{-3/2-1/2} = -\sqrt{3} Q_1 / \hbar \omega_0, \\ a_{3/2-1/2} &= a_{1/2-3/2} = -a_{-1/2 \ 3/2} = -a_{-3/2 \ 1/2} = \sqrt{3} Q_2 / 2 \hbar \omega_0. \end{aligned} \quad (4 \cdot 26)$$

When a r.f. field is applied, transitions between any two levels are possible. If we write the direction cosines of the r.f. polarization with respect to a , b , c axes as $(\lambda_1, \lambda_2, \lambda_3)$, the relative intensity of the transition $m \rightarrow m'$ will be given by

$$\begin{aligned} \mathcal{I}_{mm'} &= |E_m - E_{m'}| \left| \sum_{m''} \left[\sqrt{(I - m'')(I + m'' + 1)} (a_{mm''}^* a_{m'm''} + a_{mm''}^* a_{m'm''+1}) \right. \right. \\ &\quad \times [\lambda_2 + i \{ \lambda_1 \sin(\alpha - \beta) - \lambda_3 \cos(\alpha - \beta) \}] / 2 \\ &\quad \left. \left. + a_{mm''}^* a_{m'm''} m'' [\lambda_1 \cos(\alpha - \beta) + \lambda_3 \sin(\alpha - \beta)] \right] \right|^2. \end{aligned} \quad (4 \cdot 27)$$

The secular broadening and the thermal transition probability corresponding to this line are given by

$$(1/T_2')_{mm'} = (A_1^2 \tau_0' / 4 \hbar^2) \left\{ \sum_{m''} m''^2 (|a_{mm''}|^2 - |a_{m'm''}|^2) \right\}^2, \quad (4 \cdot 28)$$

$$W_{mm'} = (A_1^2 \tau_0' / 2\hbar^2) \left| \sum_{m''} m'' a_{mm''}^* a_{m'm''} \right|^2. \quad (4 \cdot 29)$$

If we content ourselves with the perturbation calculation, we get from (4·26)

$$\begin{aligned} (1/T_2')_{3/2 \ 1/2} &= (1/T_2')_{-3/2 \ -1/2} = (A_1^2 \tau_0' / 4\hbar^2) [1 - 3(Q_1/\hbar\omega_0)^2 + 3(Q_2/2\hbar\omega_0)^2]^2, \\ (1/T_2')_{3/2 \ -1/2} &= (1/T_2')_{-3/2 \ 1/2} = (A_1^2 \tau_0' / 4\hbar^2) [1 + 3(Q_1/\hbar\omega_0)^2 - 3(Q_2/2\hbar\omega_0)^2]^2, \quad (4 \cdot 30) \\ (1/T_2')_{1/2 \ -1/2} &= (A_1^2 \tau_0' / 4\hbar^2) [1 + 9(Q_1/\hbar\omega_0)^2 - 9(Q_2/2\hbar\omega_0)^2]^2, \\ (1/T_2')_{3/2 \ -3/2} &= (9A_1^2 \tau_0' / 4\hbar^2) [1 + (Q_1/\hbar\omega_0)^2 - (Q_2/2\hbar\omega_0)^2]^2, \\ W_{3/2 \ 1/2} &= W_{-3/2 \ -1/2} = (A_1^2 \tau_0' / 2\hbar^2) \cdot 12(Q_1/\hbar\omega_0)^2, \\ W_{3/2 \ -1/2} &= W_{-3/2 \ 1/2} = (A_1^2 \tau_0' / 2\hbar^2) \cdot 3(Q_2/2\hbar\omega_0)^2, \quad (4 \cdot 31) \\ W_{1/2 \ -1/2} &= (A_1^2 \tau_0' / 2\hbar^2) \cdot 0(Q/\hbar\omega_0)^6, \\ W_{3/2 \ -3/2} &= (A_1^2 \tau_0' / 2\hbar^2) \cdot 0(Q/\hbar\omega_0)^6. \end{aligned}$$

We see that when the quadrupole interaction is small as compared with the hyperfine interaction, the relation $1/T_2 \cdot W$ holds. Moreover, the transition probabilities corresponding to $1/2 \leftrightarrow -1/2$ and $3/2 \leftrightarrow -3/2$ are far smaller than the others so that the bypath relaxations $1/2 \leftrightarrow \pm 3/2 \leftrightarrow -1/2$ and $3/2 \leftrightarrow \pm 1/2 \leftrightarrow -3/2$ will overcome the direct ones. Any way, the relaxation in this case depends primarily on the quadrupole interaction.

According to the paramagnetic resonance data on certain diluted crystals⁵⁹ we have for copper $P/A \sim 10^{-1}$ and accordingly $PI^2/AIS \sim 1/3$. Though the values of P will be different from this in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, the magnitude will be of the same order. Therefore, the perturbation calculation may not be very good. Experiment will make this point clearer.

The line widths of the absorption lines will be of the same order of magnitude as those calculated in § 3 whatever the quadrupole interaction may be, while the relative magnitudes of the line widths will depend on the quadrupole interaction.

§ 5. Discussions

The arguments given in the preceding sections are based, below the Curie point, on the spin wave approximation which is considered to be very good at sufficiently low temperatures. In fact, for T_1 of protons in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ this approximation gave a result which agreed well with the experiment. Therefore, the predictions given above seem to be likewise reliable.

We expect that the nuclear resonance of the transition elements in antiferromagnetic substances will be detectable at very low temperatures. This is the inference drawn from the understanding that only the low frequency part of the fluctuating magnetic field can contribute to the line width and to the relaxation. The low frequency components of the frequency spectra are strongly dependent on temperature and they decrease markedly as temperature is lowered.

We neglected the low frequency components of $\langle \{\partial S_+(t) \partial S_-\} \rangle$. As we stated before, this neglect is not strictly correct because of the effect of collisions between spin waves or the effect of the higher order terms in the Hamiltonian. These terms, if any, will make much smaller contributions to the relaxation than $\langle \{\partial S_z(t) \partial S_z\} \rangle$ does.

We may say that above the Curie point the line width $1/T_2$ and the relaxation rate $1/T_1$ are of the same order of magnitude, while below the Curie point $1/T_2$ and $1/T_1$ are not in general of the same order of magnitude but the relation $1/T_2 \gtrsim 1/T_1$ holds generally. $1/T_2$ depends not so much on the direction of the local magnetic field or the quadrupole interaction. On the other hand $1/T_1$ depends sensitively on them.

As for the resonance frequency, we can estimate their order of magnitude from (3.2). They are ~ 700 Mc for Mn^{++} , ~ 1300 Mc for Co^{++} and ~ 150 Mc for Cu^{++} . These values indicate only the order of magnitude. The actual absorption lines are split by the quadrupole interactions. These frequencies are functions of the sublattice magnetization and depend on temperature.

We can expect from this type of experiments much interesting information on antiferromagnetism; for example, the temperature dependence of the sublattice magnetization, the nature of the fluctuations of the electronic spins, properties of the ground state wave functions in a crystalline field, the magnetic anisotropies, etc.

Another interesting experiment will be the resonance on such nucleus that lies neighboring to the magnetic ions. Recently, several authors^{(8), (9), (10), (11)} have stressed the partial covalent character of ionic crystals. If there are any covalency between the considering nucleus and the magnetic ions, the nuclear resonance will show a large quadrupole splitting and a large shift below T_c due to the hyperfine interaction. Examples will be F in MnF_2 discussed in I, Cl in $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, etc.

We expect further that the nuclear resonance in ferromagnetic metals such as Co, Gd, will also be detectable at sufficiently low temperatures. This will be possible even in the absence of the external magnetic field, and accordingly powder samples may be used.

It is a pleasure for the author to thank Professor T. Nagamiya at Osaka University for his encouragement and advice, and Associate Professor M. Shimizu for valuable discussions.

Appendix

We wish to derive (2.22)–(2.25) from (2.7)–(2.9). This can be reduced to the calculation of the following expression:

$$\chi(t) = \int_0^t d\tau (t-\tau) \langle \{\partial S(\tau) \partial S\} \rangle. \quad (\text{A} \cdot 1)$$

If we divide the frequency spectra into two components, i.e. the high and the low frequency parts, the correlation function can be written as

$$\langle \{\partial S(\tau) \partial S\} \rangle = \int_{-\infty}^{\infty} e^{i\omega\tau} g_L(\omega) d\omega + \int_{-\infty}^{\infty} e^{i\omega\tau} g_H(\omega) d\omega. \quad (\text{A} \cdot 2)$$

$g_L(\omega)$ and $g_H(\omega)$ are schematically drawn in Fig. 2. The first term in (A.2) will be a function which decays rapidly as time elapses. $g_H(\omega)$ is a function which equals zero except in a region between $2\omega_{\min}$ and $2\omega_{\max}$, ω_{\min} and ω_{\max} being respectively the minimum and the maximum frequencies of the spin wave spectra. Thus (A.1) becomes

$$\chi(t) = |t| \int_0^\infty d\tau \int_{-\infty}^\infty e^{i\omega\tau} g_L(\omega) d\omega + 2 \int_{2\omega_{\min}}^{2\omega_{\max}} d\omega g_H(\omega) \frac{1 - \cos \omega t}{\omega^2}. \quad (\text{A.3})$$

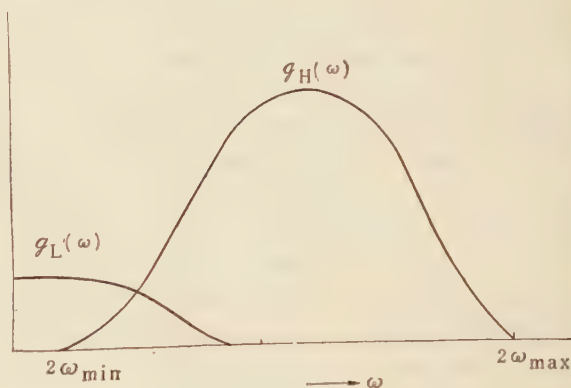


Fig. 2.

This reduces to

$$\chi(t) = \pi g_L(0) |t| - 2 \frac{g_H(c)}{c} \frac{\sin \eta t - \sin \xi t}{t}, \quad (\text{A.4})$$

where the shape of $g_H(\omega)$ is assumed to have a one maximum at $\omega=c$, and ξ and η satisfy $2\omega_{\min} \leq \xi \leq c$ and $c \leq \eta \leq 2\omega_{\max}$. The second term of (A.4) is estimated as

$$\sim (1/\omega_c^2) \cdot (1/\omega_c t) \quad \text{for } \omega_c t \gg 1.$$

This is a very small quantity and is naturally neglected. This is also the case when $g_H(\omega)$ has a finite number of maximums. We have

$$g_L(0) = \frac{1}{\pi} \int_0^\infty \langle \{\delta S(\tau) \delta S\} \rangle d\tau,$$

so that

$$\chi(t) = |t| \int_0^\infty \langle \{\delta S(\tau) \delta S\} \rangle d\tau.$$

We get (2.22) – (2.25) in the same way.

References

- 1) T. Moriya, Phys. Rev. **101** (1956), 1435; Prog. Theor. Phys. **16** (1956), 33.
- 2) R. K. Wangsness and F. Bloch, Phys. Rev. **89** (1953), 728.
- 3) R. Kubo and K. Tomita, J. Phys. Soc. Japan **9** (1954), 888.
- 4) P. W. Anderson and P. R. Weiss, Revs. Mod. Phys. **25** (1953), 269.
- 5) K. D. Bowers and J. Owen, Rep. Prog. Phys. **18** (1955), 304.
- 6) G. E. G. Hardeman, Conférence de physique des basses températures, Paris, Septembre, 1955; Hardeman, Poulis and Van Der Lugt, Physica **22** (1956), 48.
- 7) T. Nagamiya, K. Yosida and R. Kubo, Adv. Phys. **4** (1955), 1.
- 8) T. Kanda, J. Phys. Soc. Japan **10** (1955), 85.
- 9) K. Yosida and T. Moriya, J. Phys. Soc. Japan **11** (1956), 33.
- 10) J. Owen, Proc. Roy. Soc. **A227** (1955), 133.
- 11) M. Tinkham, Phys. Rev. **100** (1955), 1792; Proc. Roy. Soc. **A 236** (1956), 535, 549.

Note added in proof

Recently Shulman and Jaccarino [Phys. Rev. **103** (1956), 1126] reported a measurement of the fluorine nuclear resonance in MnF_2 above the Curie point. They observed a large shift of the resonance field which is attributed to the effect of the transfer of an electron of closed fluorine shell to the Mn^{++} ions. This electron transfer was predicted in the preceding paper I. The relaxation time of F nuclei in this case both above and below the Curie point may be calculated by using the formulas in § 2 of the text, taking the following interaction Hamiltonian:

$$\mathfrak{H} = (S + \frac{1}{2})^{-1} \sum_i (\lambda_{si} a \mathbf{S}_i \cdot \mathbf{I} + \lambda_{pi} \mathbf{S}_i \cdot \mathbf{A}_i \cdot \mathbf{I})$$

where

$$a = (8\pi/3) \gamma_e \gamma_N \hbar^2 |\psi_{2s}(0)|^2,$$

$$\mathfrak{H}_i = \gamma_e \gamma_N \hbar^2 \int |\psi_{2p}(r)|^2 r^{-5} (3\mathbf{r}\mathbf{r} - r^2) dv \quad (\text{dyadic}),$$

$\psi_{2s}(r)$, $\psi_{2p}(r)$ being the 2s- and 2p-functions of an F^- ion; λ_{si} , λ_{pi} are the transfer rates of 2s and 2p electrons of the F^- ion to the i -th Mn^{++} ion, \mathbf{S}_i the spin of the i -th Mn^{++} ion and \mathbf{I} the nuclear spin of the considering fluorine ion.

If we take only the isotropic part of the hyperfine interaction we get

$$\begin{aligned} (T_1)_{\text{h.f.}} / (T_1)_{\text{dip}} &= 9 \sum_i \langle r_i^{-6} \{F_i(\alpha\beta\gamma) + F_i'(\alpha\beta\gamma)\} \rangle / 2(8\pi/3)^2 |\psi_{2s}(0)|^4 \sum_i \lambda_{si}^2 \\ &\cong 1.6 \times 10^{-6} / \sum_i \lambda_{si}^2 \quad (T_1 = T_2) \end{aligned}$$

above the Curie point. As is seen in § 2, below the Curie point T_1 and T_2 will depend considerably on the direction of the average magnetic field at the nucleus and accordingly on the direction of the external field. When the external field is parallel to the c -axis and is smaller than the critical field, we get

$$\begin{aligned} (T_2)_{\text{h.f.}} / (T_1)_{\text{dip}} &= 9 \sum_i \langle r_i^{-6} F_i(\alpha\beta\gamma) \rangle / (8\pi/3)^2 |\psi_{2s}(0)|^4 \sum_i \lambda_{si}^2 \\ &= 1.1 \times 10^{-6} / \sum_i \lambda_{si}^2 \end{aligned}$$

at temperatures far below the Curie point.

Estimating λ from Shulman Jaccarino's data we get $(T_1)_{\text{h.f.}} / (T_1)_{\text{dip}} \sim 10^{-1}$ and therefore $(T_1)_{\text{h.f.}} \sim 10^6$ sec above the Curie point in agreement with the experimental data [(T_1)_{dip} was calculated in I]. Our preceding order estimation (given in page 41 of I) was not sound, i. e. the ratio of the average values of the local dipolar and hyperfine fields should be replaced by the ratio of the amplitudes of their fluctuations

Two-Nucleon Problem with Pion Theoretical Potential, IV —Photodisintegration of Deuteron at 20 Mev—

Junji IWADARE,* Shoichiro OTSUKI,** Mitsuo SANO,*** Shuji TAKAGI†
and Wataro WATARI††

*Research Institute for Fundamental Physics, Kyoto University, Kyoto,

**Department of Physics, Kyoto University, Kyoto,

***Institute for Nuclear Study, University of Tokyo, Tokyo,

†Department of Physics, Osaka University, Osaka,

††Department of Physics, Hiroshima University, Hiroshima

(Received August 29, 1956)

In the present paper we discuss the photodisintegration of the deuteron at about 20 Mev assuming the pion theoretical potential to make a check for the validity of the potential. A detailed calculation shows that the pion theoretical potential gives a qualitative agreement with experiments, especially in predicting a large isotropic part in the angular distribution of the final nucleons. This agreement is mainly attributed to the strong tensor force in the triplet P states and the possible hard-core-like repulsion in the singlet S state which was found in the previous analyses phenomenologically. It is to be noted, however, that there remain some ambiguities concerning the effects of the interaction moment at small inter-nucleon distances.

§ 1. Introductory remarks

In a series of works under the same title,¹⁾ we examined the nuclear forces in two-nucleon system from the point of view of the pion theory. There we established the following four points: 1) The one-pion-exchange potential in the outer region I ($r > 1.5 \times b/\mu c$ (μ : the pion mass)) predicted by the pion theory plays an essential role in the two-nucleon problem at low energies. 2) The effective coupling constant $g_c^2/4\pi$ of this potential must lie in a narrow region $g_c^2/4\pi = 0.08 \pm 0.01$ in order to give the quantitative fit to the experimental values of the low energy parameters. 3) In the region II ($1.5 \times b/\mu c > r > 0.7 \times b/\mu c$), pion theoretical corrections to the one-pion-exchange potential (the two-pion-exchange potential and so forth) possess the qualitative properties required by the experimental data. 4) This pion theoretical potential, together with suitable phenomenological parameters standing for the unknown interactions at small distances (the region III $r < 0.7 \times b/\mu c$) can explain consistently the scattering experiments at the low and intermediate energy regions (< 20 Mev). In addition to these, we obtained some indication on the features of the interactions in the region III.

The photodisintegration of the deuteron as well as the radiative capture of the neutron by the proton has served as a source of our knowledge of the interactions in the neutron-proton system. However, if one wants to draw some useful information on the shape

of the nuclear potential out of these experiments, one will find that energy regions available for the purpose are rather restricted. On the one hand, the experimental data in the low energy region $E_\gamma < 10$ Mev are simply explicable in terms of the shape independent approximation (i. e. by the effective range theory) and no information on the shape of the potential can be expected from them.⁽¹⁾ On the other hand, the situation becomes very complicated at energies higher than about one hundred Mev. The higher multipole radiation and hence the higher waves in the final state contribute to the transition. Furthermore, at such high energies the nuclear interaction in the final state might not be represented by the concept of potential. Also various other mesonic effects including the interaction moment may play an important role. In fact, the experiments at energies higher than 100 Mev⁽²⁾ show many peculiar features such as the rapid rise of the excitation function and the very large isotropic component of emitted particles, and we have not any reasonable interpretation for them. Therefore, the experiments in the intermediate energy region $10 \text{ Mev} < E_\gamma < 50 \text{ Mev}$ are most suitable for the purpose. In the present paper, the photodisintegration of the deuteron at about 20 Mev is discussed. In the discussion, the conclusions of the preceding papers of this series are taken into account.

§ 2. Photodisintegration of the deuteron at intermediate energies

There have been two experiments at about 20 Mev in which the angular distribution as well as the total cross section are measured.⁽¹⁾ The angular distribution is expressed in the form

$$\sigma(\theta) = a + b \sin^2 \theta (1 + 2\beta \cos \theta).$$

The values of the parameters a/b and β , and the total cross section are listed in Table I. The ratio a/b is measured as about 0.1 and the parameter β of the forward-backward asymmetry also amounts to 0.1.

Table I

| | detection | E_γ | a/b | β | σ_{tot} |
|------------------------------|-------------|-------------|-----------------|---------|----------------------------|
| Halpern and Weinstock (1953) | emulsion | 18 ~ 22 Mev | 0.13 ± 0.04 | 0.09 | $0.54 \pm 0.11 \text{ mb}$ |
| Allen (1955) | ZnS counter | ~ 20 Mev | 0.09 ± 0.01 | 0.1 | $0.51 \pm 0.16 \text{ mb}$ |

The measurement of the total cross section is rather uncertain because of the difficulty in monitoring the γ -rays. The experiments show this is about 0.5 mb.

A few attempts to interpret the experiments taking into account the P -state interaction were since performed. Austern⁽³⁾ calculated the isotropic ratio a/b in the energy region $15 \text{ Mev} < E_\gamma < 90 \text{ Mev}$ with a weak tensor force or a singular $\mathbf{L} \cdot \mathbf{S}$ force in the triplet P -states. The D -state component of the deuteron was also included. Berger⁽⁴⁾ also calculated a/b for the γ -ray energy $6 \text{ Mev} < E_\gamma < 28 \text{ Mev}$, assuming three phenomenological potentials. Their results were that the ratio a/b was so small compared

with the experimental values, and that the inclusion of the plausible interaction moment could not improve the situation. In the works they used the deuteron wave functions whose D -state probabilities are small compared with that of the pion theoretical one. Their use of the Born approximation in estimating the effect of the triplet P -state interaction is also hardly justifiable from our standpoint.

The effects of the interaction moment were investigated by Yoshida and Sugie⁷⁾ who determined it so as to fit the magnetic moment of ^3H and ^3He . In the absence of the triplet P -interaction, the interaction moment with radial dependence of Yukawa or exponential shape was not able to give the large a/b of 0.1 for reasonable values of the parameters. The interaction moment of the Villars type with the pion-nucleon coupling constant $g^2/4\pi=0.33$ required to explain the magnetic moment of ^3H and ^3He gave $a/b=0.033$, which is only $1/3$ of the experimental value. This coupling constant is, moreover, too large in the light of the present pion theory.

In such a circumstance, it is very interesting to investigate whether the pion theoretical potential is able to reproduce the experimental data of the photodisintegration of the deuteron, in particular, the unexpectedly large isotropic ratio. The pion theoretical potential, whose validity was established in the analyses of nucleon-nucleon scattering, is expected to survive under this criterion. In the following paragraphs we shall show that it is really the case.

§ 3. Calculation

The differential cross section of the photodisintegration of the deuteron is given as⁸⁾

$$d\sigma = \frac{e^2}{\hbar c} \frac{\hbar\omega}{Mc^2} \left(\frac{Mc}{b} \right)^2 \frac{1}{3k} \sum_{m_0} \frac{1}{2} \sum_i \sum_m |\Psi_i^m, \mathbf{O}_e \Psi_i^{m_0}|^2 \frac{d\Omega}{4\pi}. \quad (1)$$

\mathbf{O}_e is the effective interaction operator representing the absorption of a photon of momentum ω/c and polarization \mathbf{e} . Its explicit form is read as

$$\mathbf{O}_e = \frac{1}{2} \mathbf{e} \cdot \mathbf{r} + \frac{\hbar}{2Mc} \boldsymbol{\kappa} \times \mathbf{e} \cdot \mathbf{M} + i \frac{\omega}{8c} \mathbf{e} \cdot \mathbf{r} \boldsymbol{\kappa} \cdot \mathbf{r} + \dots,$$

The notations here used are all similar to those of reference 8. The first and the third terms give the electric dipole ($E1$) and quadrupole ($E2$) transitions respectively and the second one gives rise to the magnetic dipole ($M1$) transition. \mathbf{M} is the magnetic dipole operator in the unit of the nuclear magneton.

$$\begin{aligned} \mathbf{M} &= \sum_{i=1,2} (\sigma^{(i)}/2) \{ \mu_n(1 - \tau_3^{(i)}) + \mu_p(1 + \tau_3^{(i)}) \} + \frac{1}{2} \mathbf{L} + \mathbf{I} \\ &= \{ (\mu_n + \mu_p) + \frac{1}{2} (\mu_p - \mu_n) (\tau_3^{(1)} + \tau_3^{(2)}) \} \mathbf{S} + \frac{1}{4} (\mu_p - \mu_n) (\sigma^{(1)} - \sigma^{(2)}) (\tau_3^{(1)} - \tau_3^{(2)}) \\ &\quad + \frac{1}{2} \mathbf{L} + \mathbf{I}, \end{aligned}$$

where $\mathbf{S} = \frac{1}{2} (\sigma^{(1)} + \sigma^{(2)})$ and $\mathbf{L} = i\mathbf{r} \times \nabla$.

As to the interaction moment operator \mathbf{I} , we use the moment of the Villars type which

is given by the one-pion-exchange process.¹¹⁾ As in the case of the one-pion-exchange potential this Villars' moment may be reliable in the outer region. The correct shape of the interaction moment at small inter-nucleon distances cannot be given by the pion theory at the present stage. Therefore we adopt the interaction moment of the Villars' type only in the outer region $r > 1 \times \hbar/\mu c$ and cut it off to be zero in the inner region $r < 1 \times \hbar/\mu c$ in the present paper (μ : the pion mass). This procedure introduces inevitably a large uncertainty in the final results, but it serves to estimate the effect of the interaction moment. The unknown effect from the small inter-nucleon distances should be included in the $M1$ transition in the future theory.

$$\mathbf{I} = (\boldsymbol{\tau}^{(1)} \times \boldsymbol{\tau}^{(2)}) \left[(\boldsymbol{\sigma}^{(1)} \times \boldsymbol{\sigma}^{(2)}) f(x) + \left\{ 3 \frac{(\mathbf{r} \cdot \boldsymbol{\sigma}^{(1)} \times \boldsymbol{\sigma}^{(2)}) \mathbf{r}}{r^2} - (\boldsymbol{\sigma}^{(1)} \times \boldsymbol{\sigma}^{(2)}) \right\} g(x) \right],$$

$$f(x) = \begin{cases} (g_c^2/4\pi) (M/\mu) \left(\frac{2}{3} - \frac{1}{3x} \right) e^{-x} & x > 1 \\ ? & x < 1, \end{cases}$$

$$g(x) = \begin{cases} -(g_c^2/4\pi) (M/\mu) \left(\frac{1}{3} + \frac{1}{3x} \right) e^{-x} & x > 1 \\ ? & x < 1, \end{cases}$$

$$(x = (\mu c/\hbar) r).$$

In Eq. (1), $\Psi_i^{m_0}$ is the wave function of the deuteron and Ψ_f^m is the wave function of the final n - p system which consists of the plane wave with the wave number k and incoming spherical wave. Both $\Psi_i^{m_0}$ and Ψ_f^m are the solutions of the nuclear Hamiltonian including the two-nucleon interaction, for which we have developed some detailed discussions in the previous papers. The deuteron wave function $\Psi_i^{m_0}$ was obtained in the paper I, which can reproduce all deuteron parameters correctly. One of the most important features of the wave function is that the D -state probability P_D is about 0.06, which is larger than the value 0.04 usually adopted. As was discussed in the paper I, the large P_D is a necessary consequence of the strong tensor force of the one-pion-exchange potential.

Main contributions to the cross section come from the $E1$ transition to the 3P -states and the $M1$ transition to the 1S -state. The transitions to the 3S -state and to the higher states are small compared to the above two transitions. The isotropic part a also comes mainly from the $E1$ and $M1$ transitions and the backward-forward asymmetry β arises from the interference between the $E1$ and $E2$ transitions.

As will be discussed in § 4 a strong tensor force in the 3P -states due to the one-pion-exchange potential, a strong repulsive force in the region III in the 1S -state discussed in paper II and the large D -wave component in the deuteron state give rise to a large isotropic part in the angular distribution. The calculated value of a/b is about 0.1, a half of which being due to the $E1$ transition while the other due to the spin-flip $M1$ transition.

The contributions of the non-spin-flip $M1$ transition, the $M2$ transition and the

interference terms are also calculated, but are found to have small effects to the final results.

The main coefficients of the angular distribution are summarized in Table II. The calculation is done at $E_\pi = 22.2$ Mev, which corresponds to 40 Mev in the nucleon-nucleon scattering in the laboratory system.

Table II. Coefficients of angular distribution (in the unit of mb/ster.)

| angular distribution | 1 | $\sin^2 \theta$ | $\sin^2 \theta \cos \theta$ | $\sin^2 \theta \cos^2 \theta$ | final states |
|--|--------------------|--------------------|-----------------------------|-------------------------------|-------------------------------|
| M1 (spin-flip) A)* (non-spin-flip) B)* | 0.000 ₁ | 0.000 | | | $^1S_0, ^1D_2$ |
| | 0.005 ₃ | 0.000 ₅ | | | $^3S_1 + ^3D_1, ^3D_2$ |
| | negligible | negligible | | | |
| E1 | 0.002 ₀ | 0.049 | | | $^3P_0, ^3P_1, ^3P_2$ |
| E2 | | | | 0.001 ₃ | $^3S_1 + ^3D_1, ^3D_2, ^3D_3$ |
| Interference E1-E2 | | | 0.004 ₉ | | |

The main terms of the angular distribution can be written in two extreme cases of A) (hard core radius in the 1S -state = 0) and B) (hard core radius in the 1S -state = $0.5 \times \hbar/\mu c$)* as

$$A) \quad d\sigma/d\Omega = \{0.04 + \sin^2 \theta (1 + 0.1 \cos \theta)\} 0.049 \text{ mb/ster.},$$

$$B) \quad d\sigma/d\Omega = \{0.15 + \sin^2 \theta (1 + 0.1 \cos \theta)\} 0.049 \text{ mb/ster.}$$

The total cross section is

$$A) \quad \sigma_{\text{tot}} = 0.45 \text{ mb,}$$

$$B) \quad \sigma_{\text{tot}} = 0.52 \text{ mb.}$$

Therefore the result is consistent with the experimental data.

§ 4. Detailed discussions

4-1. E1 transition

The cross section of the pure E1 transition is written as follows:

$$\begin{aligned}
 d\sigma(\theta) = & \frac{\pi}{3} \frac{e^2}{\hbar c} \frac{M\omega}{\hbar k} \left(\frac{\hbar}{\mu c} \right)^3 \frac{d\Omega}{4\pi} \\
 & \times \left[\left\{ \frac{1}{9} |e^{i\delta_0} I_0 - e^{i\delta_2} I_2|^2 + \frac{1}{4} |e^{i\delta_1} I_1 - e^{i\delta_3} I_3|^2 \right\} \right. \\
 & \left. + \frac{1}{2} \sin^2 \theta \left\{ \frac{1}{4} |e^{i\delta_1} I_1 + 3e^{i\delta_3} I_3|^2 - \frac{1}{15} |e^{i\delta_0} I_0 - 5e^{i\delta_2} I_2|^2 + \frac{1}{15} I_0^2 \right\} \right], \quad (2)
 \end{aligned}$$

where

* With respect to the case A) and B), see Table IV.

$$I_0 = \int_0^\infty x dx v_0 (u_d - \sqrt{2} w_d), \quad I_1 = \int_0^\infty x dx v_1 \left(u_d + \frac{1}{\sqrt{2}} w_d \right),$$

$$I_2 = \int_0^\infty x dx v_2 \left(u_d - \frac{1}{5\sqrt{2}} w_d \right),$$

and v_J denotes the standing wave solution in the 3P_J -state, δ_J being its phase shift. u_d and w_d are the radial wave functions of the S - and D -components of the deuteron normalized as $\int_0^\infty (u_d^2 + w_d^2) dx = 1$. If there is no non-central force in the 3P -states, all radial waves in the 3P -states coincide with one another, i. e., $v_0 = v_1 = v_2$ and $\delta_0 = \delta_1 = \delta_2$. Then the isotropic ratio is reduced. In our calculation it reduces to about a half of its actual value. Further, if we neglect the D -state component of the deuteron, all integrals I_J 's take the same value, and the isotropic contribution of the $E1$ transition vanishes.

For a qualitative discussion, we consider the direction of the splitting of the integrals I_J . The D -state contribution diminishes I_0 considerably and I_2 a little, and makes I_1 somewhat larger, since w_d and u_d have the same sign as inferred from the positive quadrupole moment of the deuteron. The diagonal elements of the tensor potential $V_T S_{12}$ for 3P_0 -, 3P_1 - and 3P_2 -states are $-4V_T$, $2V_T$ and $-(2/5)V_T$ respectively. The pion theoretical V_T is known to be positive in the triplet odd state, and its attractive effect in the 3P_0 -state diminishes I_0 further, and the repulsion in the P_1 -state makes I_1 larger, since the factor x in the integrands makes I_J an outer quantity and a repulsive force tends to push the wave function outwards. Thus the tensor potential in the both (triplet even and triplet odd) states cooperate to increase the isotropic contribution in the $E1$ transition. This feature is characteristic in the pion theoretical potential in contrast to the phenomenological ones such as the Serber potential.

The impact parameter b of the P -wave in the two-nucleon system is about $1.6 \times \hbar/\mu c$ at the energy corresponding to $E_T \sim 20$ Mev. Therefore the one-pion-exchange potential in the region I plays the most important role to distort the wave. On the contrary, the interaction in the region III hardly affects the wave. In the region II, if we apply various derivation methods of potential, we have potentials with almost similar shapes as can be seen from Fig. 1 of the paper III. Therefore we first assume the TMO potential with $g^2/4\pi = 0.08$ and then estimate the effects of the possible corrections to the TMO potential. In estimating the effects, we examine whether the correction is consistent with the almost isotropic angular distribution of p - p scattering, since it is very sensitive to the P -wave phase shifts. Values of the phase shifts and the integrals I_J assuming the TMO potential are shown in Table III. The result is that:

$$a_{E1} = 0.0020 \text{ mb/ster.}, \quad b_{E1} = 0.049 \text{ mb/ster.},$$

$$(a/b)_{E1} = 0.041, \quad \sigma_{\text{tot}, E1} = 0.44 \text{ mb.}$$

When some correction is added only to the central force, the isotropic ratio a/b scarcely changes since all phase factors $\exp(i\delta_J)$ as well as integrals I_J change by almost the same amount. Furthermore, the pion theory predicts that the central potential is

Table III. Values of the 3P -wave phase shifts and the dimensionless integral I_J ($J=0, 1, 2$)
 The wave functions are normalized as $\int (u_d^2 + w_d^2) dx = 1$ and $v_J \xrightarrow{r \rightarrow \infty} \sin(kr + {}^3\delta_J^1 - \pi/2)$.

| State | 3P_0 | 3P_1 | 3P_2 |
|-------------------------------|-----------|-----------|--------------------|
| Phase shifts ${}^3\delta_J^1$ | 0.296 | -0.122 | 0.033 ₂ |
| $\int x dx v_J u_d$ | 1.472 | 1.711 | 1.595 |
| $\int x dx v_J w_d$ | 0.505 | 0.462 | 0.473 |
| I_J | 0.758 | 2.038 | 1.528 |

very small compared with the tensor potential. Thus the uncertainty in the central potential is masked by that of the tensor potential.

Then the effect due to the uncertainty in the tensor potential should be examined. From Table III we can find that the contribution from the tensor potential to the 3P -wave phase shifts is such that $\delta(V_t) \simeq -0.073$, i. e., the tensor potential contributes about $-4\delta(V_t)$ to ${}^3\delta_0^1$, $+2\delta(V_t)$ to ${}^3\delta_1^1$ and $-2\delta(V_t)$ to ${}^3\delta_2^1$. As has been mentioned before, we can except that $\delta(V_t)$ is mainly due to the one-pion-exchange potential. When the coupling constant $g_e^2/4\pi$ of the one-pion-exchange potential is changed in the region $0.07 < g_e^2/4\pi < 0.09$, $|\delta(V_t)|$ decreases or increases by about 10% or a little more. Since the signs of the one- and two-pion-exchange potentials are opposite to each other in the triplet odd state, it would be sufficient to estimate the effect due to some possible correction for the tensor potential by changing $\delta(V_t)$ by $\pm 15\%$. The result is that:

$$a/b = 0.045,$$

when the tensor potential becomes more strongly positive. When it decreases by 15%

$$a/b = 0.037.$$

Some authors derived the spin-orbit force from the pion theory. Unfortunately, their results do not coincide one another even with respect to the sign. However, the $L \cdot S$ potential may not be important at $E_\pi \sim 20$ Mev since the $L \cdot S$ potential has the range of one half of the pion Compton wave length. Thus we assume that the contribution from the $L \cdot S$ potential $|\delta(V_{LS})|$ to the P -wave phase shifts is $|\delta(V_{LS})| = 0.2 \times |\delta(V_T)|$. The $L \cdot S$ potential contributes $2\delta(V_{LS})$, $\delta(V_{LS})$ and $-\delta(V_{LS})$ to ${}^3\delta_0^1$, ${}^3\delta_1^1$ and ${}^3\delta_2^1$ respectively. In such cases, estimates show that

$$a/b = 0.042 \quad \text{for } V_{LS} > 0,$$

and
$$a/b = 0.040 \quad \text{for } V_{LS} < 0.$$

Thus, a/b changes only little, since, in the isotropic part in Eq. (2), the first term increases (decreases) and the second decreases (increases) when the $L \cdot S$ potential is taken into account.

4.2. $M1$ -transition (spin-flip transition)

The spin-flip cross section from the $M1$ transition is

$$d\sigma(\theta) = \frac{\pi}{3} \frac{e^2}{\hbar c} \frac{\hbar \omega}{Mc^2} \frac{1}{k} \frac{\hbar}{\mu c} \left[|J_{00} - \frac{1}{\sqrt{10}} J_{20}|^2 + \frac{1}{8} (2 + 3 \sin^2 \theta) |\sqrt{5} J_{02} + \sqrt{2} J_{22}|^2 \right. \\ \left. + \frac{1}{4} (2 - 3 \sin^2 \theta) \{ |J_{00} - \sqrt{2} J_{22}|^2 - |J_{00} + \sqrt{5} J_{02}|^2 + |J_{20} - \sqrt{5} J_{02}|^2 \right. \\ \left. - |J_{20} + \sqrt{2} J_{22}|^2 \} \right] \frac{d\Omega}{4\pi},$$

$$J_{00} = e^{i\delta_0} \int u_d \{ (\mu_p - \mu_n) - 4f(x) \} U dx, \quad J_{22} = e^{i\delta_2} \int w_d \{ (\mu_p - \mu_n) - 4f(x) - 4g(x) \} W dx,$$

$$J_{02} = e^{i\delta_2} \int u_d \cdot 4g(x) W dx, \quad J_{20} = e^{i\delta_0} \int w_d \cdot 4g(x) U dx,$$

where U and W are the radial function of the 1S - and 1D -waves, δ_0 and δ_2 being their phase shifts. The terms which contain $f(x)$ or $g(x)$ represent the contribution from the interaction moment. J_{00} is the dominant 3S - 1S transition and has a large uncertainty arising from the unknown nuclear interaction at small distances in the 1S -state as well as the interaction moment.

However, we have obtained some information on the inner interaction in the 1S_0 -state in the preceding paper, II. Namely when the one-pion-exchange potential is present in the outer region, there must be a strong attraction in the region between the pion range and a half of it, which is consistent with the prediction by the two-pion-exchange potential, and there must be a strong repulsion at small distances at least up to about $r \sim 0.2 \times \hbar/\mu c$ that pushes the wave function outside. We have also obtained a guess that the maximum core radius is about $0.6 \times \hbar/\mu c$.

It must be noticed that the matrix element of the 3S_1 - 1S_0 transition is quite sensitive to the core radius in the 1S_0 -state. The reason is that the integrand of J_{00} changes its sign at the node of the 1S_0 -wave. At 40 Mev, when the hard core is not present, the node lies at about $r \sim 2.5 \times \hbar/\mu c$ which is small compared with the deuteron radius $3.05 \times \hbar/\mu c$. In this case the contribution to J_{00} coming from outside the node and that from inside it have the comparable order of magnitude to cancel one another. Thus the matrix element almost vanishes. If the hard core is present, the inner integral is reduced and in consequence the positive contribution from the outer part remains.

We estimate the effect of the hard core on the 3S_1 - 1S_0 transition using the 1S_0 -wave function in hand, which was solved under the influence of the TMO potential. The results are shown in Table IV. The lower case, B), of the Table IV is not reasonable since the 1S_0 -wave phase shift 0.05 is too small to be consistent with the p - p scattering

Table IV. The $M1$ -transition (3S_1 - 1S_0)

| a (mb/ster.) | a/b | TMO wave function | | | |
|----------------|-------|------------------------------|------------|-----------------------------|--------------------------|
| | | core radius in $\hbar/\mu c$ | $g^2/4\pi$ | $r_e (10^{-13} \text{ cm})$ | ${}^1\delta_0$ at 40 Mev |
| A) 0.0001 | 0.002 | 0 | 0.08 | 2.04 | 0.92 |
| B) 0.0053 | 0.11 | 0.5 | 0.12 | 2.81 | 0.05 |

experiment. On the other hand, ${}^1\hat{o}_0$ can not be larger than 0.8 since otherwise the p - p scattering cross section in the singlet state alone is over the experimental data.

The most important points with respect to the 3S_1 - 1S_0 transition is that the matrix element is quite sensitive to the core radius and it may be possible to give the isotropic ratio a/b of about 0.05 without contradiction with the singlet low energy data and the p - p scattering at 40 Mev, if the pion theoretical potential has an approximate core radius of about $0.3 \times b/\mu c$. The calculation assuming the pion theoretical potential in the 1S_0 -state as discussed in the paper II is not made here, since it is clear that the qualitative conclusions obtained above assuming the TMO potential is not modified at all.

§ 5. Conclusions

In conclusion we summarize the results. At the energy region which we treat, mainly $E1$ - and $M1$ -transitions contribute to the cross section mainly. A strong tensor force in the 3P -states appearing in the one-pion-exchange potential favours to increase the isotropic part in the $E1$ transition. The matrix element of the $M1$ -transition is sensitive to the radius of the hard core in the 1S_0 -state whose necessity has been inferred from the analyses of low energy scattering experiments. A reasonable choice of the hard core radius gives the good value to the matrix element of the $M1$ -transition. As to the interaction moment and to the possible $L \cdot S$ type potential our estimate shows that they give rather small contributions to the result, though as to the former contribution a rather large uncertainty is expected to come from the small inter-nucleon distances.

Thus we can conclude that all the characteristic features of the one-pion-exchange potential favour to reproduce experimental data of the photodisintegration of the deuteron at intermediate energies, in particular, the large isotropic ratio.

References

- 1) J. Iwadare, S. Otsuki, R. Tamagaki and W. Watari, Prog. Theor. Phys. **16** (1956), 455, 472, 604.
- 2) H. A. Bethe and C. Longmire, Phys. Rev. **77** (1950), 647.
- 3) For example E. A. Whalin, B. D. Schriever and A. O. Hanson, Phys. Rev. **101** (1956), 377.
- 4) J. Halpern and E. V. Weinstock, Phys. Rev. **91** (1953), 934.
L. Allen, Phys. Rev. **98** (1955), 705.
- 5) N. Austern, Phys. Rev. **85** (1952), 283.
- 6) J. M. Berger, Phys. Rev. **94** (1954), 1698.
- 7) A. Sugie and S. Yoshida, Prog. Theor. Phys. **10** (1953), 236.
- 8) W. Rarita and J. Schwinger, Phys. Rev. **59** (1941), 556.
- 9) E. Villars, Phys. Rev. **72** (1947), 256.

On a Theory of a Composite Model of Elementary Particles

Ziro MAKI

Institute of Theoretical Physics, Nagoya University, Nagoya

(Received August 24, 1956)

It is shown that the composite model of elementary particles proposed by Sakata can be formulated as a covariant theory of bound states. The forces among fundamental particles (nucleons and Λ -particles) from which the composite pions, heavy mesons and hyperons are supposed to be constructed are derived from the assumed basic interactions of Fermi type. Equations for bound states are solved by the lowest order approximation (the chain approximation) with the cut-off device.

§ 1. Introduction and summary

Ever since de Broglie had proposed a theory of fusion for light quanta,¹⁾ several attempts concerning the compound nature of some of the elementary particles have been suggested by various authors. These approaches were restricted to be of more or less academic nature on account of the lack of decisive experimental grounds for, or against, such theories. However, the recent developments of high energy experiments which have revealed the existence of many kinds of unstable heavy particles with their interesting properties on the one hand, and a remarkable success of systematization program of these new particles developed by Nishijima and Gell-Mann^{2),3)} on the other, brought us into a new situation. That is, the compound hypothesis for heavy particles turns out to be a prominent method to explore the foundation of the success of the theory of Nishijima and Gell-Mann.

The purpose of this paper is to show a possibility of describing the composite model as a theory of bound states based on the conventional scheme of the quantum field theory. If our approach provides an adequate description of the composite model, it will become clear that the compound hypothesis is conceivable even within the framework of the current theory. On the contrary, when we fail entirely to account for several basic assumptions implied in composite models, e. g. existence of the strong and short range force between the constituent particles and so on, ultimate validity of the compound hypotheses must be investigated in connection with the deeper knowledge of the structure of the elementary particles.

In this paper, we shall discuss in detail the composite model of heavy particles proposed recently by Sakata.⁴⁾ First, we construct a theory of composite pions as a relativistic two body problem based on a direct Fermi interaction (§§ 2-4). This is essentially a hole theoretical improvement on Fermi and Yang's theory.⁵⁾ Then with the introduction of the Λ -particle as one of the fundamental particles (Sakata model) we deal

with the composite model for heavy mesons (§ 5). It will be shown that the Fermi-Yang-Sakata model can be formulated as a covariant theory of bound states under a suitable approximation method, viz. the covariant new Tamm Dancoff approximation which will be explained in § 1. In the course of calculation, the strong divergences inherent in the basic interactions are avoided by the cut-off prescription, hence the main features of our results are characterized by an effective interaction constant $g\lambda^2$ (§ 4).

The mathematical structure of our theory of bound states is quite similar to that of the recent work of Heisenberg et al.⁽¹⁾ In other words, we intend here to extract a compound hypothesis hidden in the Heisenberg theory along the line of the Sakata model by forsaking the nonlinearity problems.

It should be stressed that the remarkable success of the recent meson theory⁽⁷⁾ also provides a crucial test to the composite model. Although it was suggested by Hayakawa⁽³⁾ that the composite pion theory would not at all be inferior to the ordinary theory, a detailed analysis seems to be necessary by using some profitable mathematical scheme.

§ 2. General formalism

A. Derivation of the fundamental equations

In this section we shall explain our theory taking the composite model for the pion as an example. In order to get the pion as a bound state of the nucleon-antinucleon system, it is necessary to introduce a short range force which acts between them. According to Fermi and Yang we assume that these forces are derivable from the elementary interaction,

$$N \text{ (nucleon)} + \bar{N} \text{ (anti-nucleon)} \rightleftharpoons N + \bar{N}.$$

On account of the isotopic independence of the whole system, the interaction Lagrangian will be a linear combination of

$$: (\bar{\psi} \tau O^A \psi) (\bar{\psi} \tau O^A \psi) : \quad (2.1)$$

and

$$: (\bar{\psi} \tau_4 O^A \psi) (\bar{\psi} \tau_4 O^A \psi) : , \quad (2.2)$$

where $\psi = \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \\ \psi^{(3)} \\ \psi^{(4)} \end{pmatrix}$ denotes the nucleon field operator, $\tau (\tau_1, \tau_2, \tau_3)$ and τ_4 the isotopic spin of the nucleon, and $O^A (A=1, 2, 3, 4, 5)$ are the Dirac matrices which give S (scalar), V (vector), T (tensor), A (pseudovector) and P (pseudoscalar) types of invariant Fermi interactions respectively. From the general combination of (2.1) and (2.2) we take for simplicity

$$-\mathcal{L}' = (g_A/4) \sum_{i=1}^4 : (\bar{\psi} \tau_i O^A \psi) (\bar{\psi} \tau_i O^A \psi) : \quad (2.3)$$

as the basic interaction of our compound theory. The summation over the superscript A is not necessary if we do not take a mixture of the different Fermi interactions. Then the equation of motion for ψ reads

$$(\gamma_\mu \partial_\mu + \kappa) \psi(x) = - (g_A/2) \tau_i O^A : \bar{\psi}(x) \tau_i O^A \psi(x) : , \quad (2.4)$$

where κ is the nucleon mass.*

The pion in our model may be described by a covariant Fock-representation,

$$\left(\begin{array}{c} \varphi(x|\gamma) \\ \varphi(x_1, x_2|\gamma_1, \gamma_2) \\ \dots\dots\dots \\ \varphi(x_1, \dots, x_n|\gamma_1, \dots, \gamma_n) \\ \dots\dots\dots \end{array} \right), \quad (2.5)$$

in which φ is the Feynman amplitude

$$\varphi(x_1, \dots, x_n|\gamma_1, \dots, \gamma_n) = \langle 0 | N(\psi(x_1) \dots \psi(x_n) \bar{\psi}(\gamma_1) \dots \bar{\psi}(\gamma_n)) | \pi \rangle \quad (2.6)**$$

with a suitably defined N -product. For $n=1$, φ reduces to

$$\varphi(x|\gamma) = \langle 0 | T(\psi(x) \bar{\psi}(\gamma)) | \pi \rangle,$$

where $T(\dots\dots)$ denotes the chronological ordering product.

The simultaneous equations for the wave functions in (2.5) are easily obtained by the successive use of the field equation (2.4). We obtain first the equation for $\varphi(x|\gamma)$:

$$(\gamma_\mu \partial_\mu + \kappa)_x \varphi(x|\gamma) = -(g_A/2) \tau_i O^A \langle 0 | T[\psi(x) (\bar{\psi}(x) \tau_i O^A \psi(x)) : \bar{\psi}(\gamma)] | \pi \rangle, \quad (2.7)$$

which is re-expressed by decomposing T -product in the right-hand side into N -products as follows:

$$\begin{aligned} (\gamma_\mu \partial_\mu + \kappa)_x \varphi(x|\gamma) = & -(g_A/2) \tau_i O^A \langle 0 | N[\psi(x) (\bar{\psi}(x) \tau_i O^A \psi(x)) \bar{\psi}(\gamma)] | \pi \rangle \\ & - (g_A/2) \tau_i O^A \varphi(x|x) \tau_i O^A S_F(x-\gamma) \\ & + (g_A/2) \tau_i O^A \text{Tr}[\tau_i O^A \varphi(x|x)] S_F(x-\gamma). \end{aligned} \quad (2.8)$$

As the nucleon Green's function S_F , we adopt here the free nucleon propagator:

$$\left. \begin{aligned} S_F(x-\gamma) = & \frac{i}{(2\pi)^4} \int \frac{i \not{p} - \kappa}{p^2 + \kappa^2 - i\varepsilon} e^{ip(x-\gamma)} d^4 p, \\ (\gamma_\mu \partial_\mu + \kappa)_x S_F(x-\gamma) = & -i\delta(x-\gamma). \end{aligned} \right\} \quad (2.9)$$

This, in turn, implies that our N -products obey the definition of Matthews and Salam.⁽⁹⁾

In a similar way, the equations for $\varphi(x_1, x_2|\gamma_1, \gamma_2)$ are to be derived, for which we require the equations for $\varphi(x_1, x_2, x_3|\gamma_1, \gamma_2, \gamma_3)$, and so on. The set of equations thus obtained can be solved by means of the new Tamm-Dancoff (NTD) method,⁽¹⁰⁾ by imposing the condition $\varphi(x_1, \dots, x_n|\gamma_1, \dots, \gamma_n) = 0$ for $n \geq N$ (N being some large number).

Though it may be expected that the NTD method will be applicable in so far as we confine ourselves to the cases of the low energy stationary states of non-singular systems (our system will become non-singular if we introduce the cut-off hypothesis.),

* We simply regard κ as the experimental mass of the nucleon.

** $|0\rangle$ is the true vacuum and $|\pi\rangle$ the one pion state.

we are uncertain that for what number N the configurations $\varphi(x_1 \cdots x_n | y_1 \cdots y_n)$ ($n \geq N$) are safely negligible. Nevertheless, it seems to be natural to suppose that the lowest order configuration, $\varphi(x|y)$, plays more or less important part in our system, since the hole-theoretical wave function, $\varphi(x|y)$, includes at least partially the effects of infinite number of virtual nucleon-antinucleon pairs in contrast to the corresponding wave function of the ordinary Tamm-Dancoff approach, even if we neglect all φ 's with $n \geq 2$. (See Fig. 1.) In this paper we shall, as a first step, investigate our system under the lowest order approximation assuming $\varphi=0$ for $n \geq 2$. Then (2.8) becomes

$$(\gamma_\mu \partial_\mu + \kappa)_x \varphi(x|y) = (g_A/2) \tau_i O^A \text{Tr}[\tau_i O^A \varphi(x|x)] S_F(x-y) - (g_A/2) \tau_i O^A \varphi(x|x) \tau_i O^A S_F(x-y). \quad (2.10)$$

For the charged pion, say π^+ , we see

$$(\gamma_\mu \partial_\mu + \kappa)_x \varphi^{p\bar{n}}(x|y) = g_A O^A \text{Tr}[O^A \varphi^{p\bar{n}}(x|x)] S_F(x-y) \quad (2.11)$$

with

$$\varphi^{p\bar{n}}(x|y) \equiv \langle 0 | T(\psi^p(x) \bar{\psi}^{\bar{n}}(y)) | \pi^+ \rangle.$$

This can also be expressed as a Nambu equation¹¹⁾:

$$(\gamma_\mu \partial_\mu + \kappa)_x (\gamma_\mu^T \partial_\mu - \kappa)_y \varphi^{p\bar{n}}(x|y) = i g_A O^A \text{Tr}[O^A \varphi^{p\bar{n}}(x|x)] \partial(x-y). \quad (2.12)$$

The corresponding integral equation for the bound states will be

$$\varphi^{p\bar{n}}(x|y) = i g_A \int S_F(x-\hat{\xi}) O^A \text{Tr}[O^A \varphi^{p\bar{n}}(\hat{\xi}|\hat{\xi})] S_F(\hat{\xi}-y) d^4 \hat{\xi}, \quad (2.13)$$

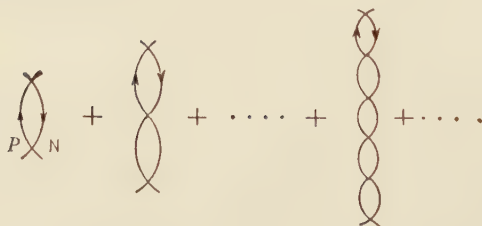


Fig. 1

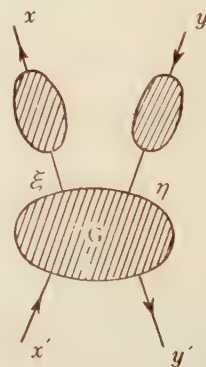


Fig. 2

which contains the interactions represented in terms of the Feynman diagrams such as Fig. 1. We may, therefore, call our approximation "the chain approximation." It should be noted that the equation (2.13) comes also from the lowest order approximation for the exact Bethe-Salpeter equation

$$\varphi^{p\bar{n}}(x|y) = \int S_{F'}(x-\hat{\xi}) G(\hat{\xi}, y'; x', y) S_{F'}(y-y) \varphi^{p\bar{n}}(x'|y') d^4 \hat{\xi} d^4 y' d^4 x' d^4 y' \quad (2.14)$$

in which the interaction kernel G (see Fig. 2.) is replaced by the first term of its expansion in g_A :

$$G \rightarrow i g_A \delta(\xi - x') \delta(\eta - y') \delta(\xi - \eta) O^A \cdot O^A.$$

It is obvious that the equation (2.13), as it stands, has no solutions because of the singularity of the interaction. Therefore we avoid the difficulty by introducing a cut-off prescription.* In order to keep the covariance of our equation, we make use of Feynman's cut-off in the momentum representation. Separating the centre of mass coordinates by

$$\left. \begin{aligned} \varphi^{\bar{m}}(x|y) &= e^{iP_\mu X_\mu} \tau(z) \\ \text{where } P_\mu &: \text{ the total energy-momentum of the system} \\ \text{and } X_\mu &= \frac{1}{2}(x+y)_\mu, \quad z_\mu = (x-y)_\mu \end{aligned} \right\},$$

we obtain, instead of (2.13), the equation

$$\tau(z) = -i \frac{g_A}{(2\pi)^4} \int S_F(P+p) O^A \text{Tr}[O^A \tau(0)] S_F(p) e^{i(P/2+p)z} d^4 p.$$

Note that $\tau(z)$ depends trivially on the relative coordinates z . The cut-off factor $C(p)$ may be attached in this representation:

$$\tau(0) = -i \frac{g_A}{(2\pi)^4} \int S_F(P+p) O^A \text{Tr}[O^A \tau(0)] S_F(p) C^2(p) d^4 p, \quad (2.15)$$

or

$$\tau(0) = -i \frac{g_A}{(2\pi)^4} \int S_F\left(p + \frac{P}{2}\right) C\left(p + \frac{P}{2}\right) O^A \text{Tr}[O^A \tau(0)] C\left(p - \frac{P}{2}\right) S_F\left(p - \frac{P}{2}\right) d^4 p. \quad (2.15')$$

$C(p)$ is defined tentatively by

$$C(p) = (-\kappa^2 + \lambda^2) / (p^2 + \lambda^2) \quad (2.16)$$

where λ represents a "cut-off mass." Although the cut-off procedure often causes contradictions with the general principles of the quantum field theory, e. g., the causality, the gauge invariance and others, we shall not make mention of this problem in this work. The space-time extension of the interaction due to the cut-off may be picturized in the "spring model interpretation" discussed by Koba.^[12] The interactions in (2.15) and (2.15') are then indicated by the diagrams Fig. 3a and



Fig. 3a



Fig. 3b

* Rigorously speaking, it is expected that two or more mass states of the nucleon may appear in our scheme as a result of the cut-off. However, since it is quite unlikely that these spurious nucleons exist with the mass close to the observable one, we leave them out of consideration. Naturally our equation (2.13) does not contain such spurious nucleon states.

Fig. 3b, where the rugged lines represent the "cut-off propagator" $C(p)$.

B. Unwanted solutions and neutral mesons

(1) It is clear that our basic interaction should be one which gives a strong attractive force between the nucleon and the anti-nucleon, and provides a repulsive or less attractive force between two nucleons. The restrictions of this kind imposed on the type of basic interactions are common in any theory of the composite particle model. The Nambu equation for $N-N$ system is found to be

$$(\gamma_\mu \partial_\mu + \kappa)_x (\gamma_\mu \partial_\mu + \kappa)_{x'} \bar{\varphi}^{p,N}(x, x') \\ = -ig_A O^A C(\bar{\varphi}^{p,N}(x', x))^T C^{-1} C O^{AT} C^{-1} \delta(x - x'), \quad (2 \cdot 17)^*$$

where

$$\bar{\varphi}^{p,N}(x, x') = \langle 0 | T(\psi^p(x), \psi^N(x')) | \rangle C^{-1}. \quad (C: \text{charge conjugation matrix})$$

Therefore, it is necessary that (2.17) has no bound states so that the present theory may be valid.

(2) The neutral mesons can be expressed by the Fock-representation:

$$\begin{bmatrix} \varphi^{p\bar{p}}(x|\gamma) \\ \varphi^{n\bar{n}}(x|\gamma) \end{bmatrix}$$

in which $\varphi^{p\bar{p}}$ and $\varphi^{n\bar{n}}$ obey the equation

$$(\gamma_\mu \partial_\mu + \kappa)_x \varphi^{p\bar{p}}(x|\gamma) = g_A O^A \{ \text{Tr}[O^A \varphi^{p\bar{p}}(x|x)] \\ - \varphi^{p\bar{p}}(x|x) O^A - \varphi^{n\bar{n}}(x|x) O^A \} \mathcal{S}_F(x - \gamma) \quad (2 \cdot 18)$$

and the same equation for $p(\bar{p}) \leftrightarrow n(\bar{n})$. It is easy to show that there exists a special solution

$$\varphi_1^{p\bar{p}}(x|\gamma) = -\varphi_1^{n\bar{n}}(x|\gamma), \quad (2 \cdot 19)$$

which satisfies

$$(\gamma_\mu \partial_\mu + \kappa)_x \varphi_1^{p\bar{p}}(x|\gamma) = g_A O^A \text{Tr}[O^A \varphi_1^{p\bar{p}}(x|x)] \mathcal{S}_F(x - \gamma) \quad (2 \cdot 20)$$

and the same equation for $\varphi_1^{n\bar{n}}$. This gives the neutral pion with isotopic spin $I=1$. That the same equations are derived for π^0 and π^\pm is evidently a general consequence of the isotopic independence of the basic interactions.

As to the neutral mesons with $I=0$ (the singlet mesons), we shall discuss these in detail in connection with the inclusion of the Λ -particle, since the singlet meson contains a configuration in which Λ and $\bar{\Lambda}$ (anti- Λ) compose a bound state (§ 6).

§ 3. Normalization of the wave function and the strength of the $\pi-N$ interaction

One of the most interesting features of the composite model would be the fact

* $(\dots)^T$ denotes a transposed matrix transposed about the spinor suffices in φ (and $\bar{\varphi}$).

that we can determine in principle the force between a composite particle (π) and its constituent (N) simultaneously with the determination of the bound level, i. e. the mass of the pion. The strength of the pion-nucleon interaction may be estimated by the relation

$$\langle p, 0 | \mathcal{L}'(x) | n, \pi^+ \rangle \sim g_A \langle p | \bar{\psi}^p O^A \psi^n(x) | n \rangle \langle 0 | : \bar{\psi}^n O^A \psi^p(x) : | \pi^+ \rangle \quad (3.1)$$

as was done by Fermi-Yang and Wentzel.¹³⁾ In this expression $\langle p, 0 |$ and $| n, \pi^+ \rangle$ denote the "one proton" and the "one neutron plus one π^+ " states respectively. It should be noted that we must use the normalized wave function φ_{π^+} in the right-hand side of (3.1) noticing $\langle 0 | : \bar{\psi}^n O^A \psi^p(x) : | \pi^+ \rangle = -\text{Tr}[O^A \varphi_{\pi^+}(x|x)]$.

The normalization of φ_{π^+} may be performed as follows. Let $j_\mu(0)$ be the current density operator at the point x_0 , then one obtains a normalization condition

$$\int_0 \langle \pi | j_\mu(0) | \pi \rangle d\sigma_\mu = Q \text{ (total charge)} = 1. \quad (3.2)^*$$

That is, we have only to determine the norm of φ_{π^+} in such a way that it satisfies the relation (3.2). In order to find the explicit form of $\langle \pi | j_\mu(0) | \pi \rangle$, let us modify \mathcal{L}' by introducing an "external field" q_μ :

$$\mathcal{L}' \rightarrow \mathcal{L}' - j_\mu q_\mu.$$

Our first task is to evaluate the wave function

$$\varphi_\mu(0; 1|2) \equiv \langle 0 | T(j_\mu(0) \psi^p(1) \bar{\psi}^n(2)) | \pi \rangle \quad (3.4)$$

by making use of the equation

$$[1 - \mathcal{J}(1, 2)_0] \varphi_\mu(0; 1|2) = i(\partial \mathcal{J}(1, 2) / \partial q_\mu(0))_0 \varphi(1|2), \quad (3.5)$$

where $(\dots)_0$ means that the limit $q_\mu \rightarrow 0$ is to be taken after differentiating with respect to q_μ , and $\mathcal{J}(1, 2)$ denotes the integral operator:

$$\begin{aligned} \mathcal{J}_{\alpha\beta; \rho\sigma}(1, 2) &= \int d^4x_3 \dots d^4x_6 S'_{F\alpha\tau}(1, 3) S'_{F\delta\beta}(4, 2) \\ &\times G_{\tau\rho\sigma\delta}(3, 6; 5, 4) \times \dots \end{aligned} \quad (3.6)$$

In accordance with the derivation of our fundamental equations (2.10-13) the force operator $\mathcal{J}(1, 2)$ can be simplified by the chain approximation:

$$\left. \begin{aligned} S'_F &\rightarrow S_F, \\ G_{\tau\rho\sigma\delta} &\rightarrow ig_A \delta(3-5) \delta(4-5) \delta(3-4) O_{\rho\sigma}^A O_{\tau\delta}^A. \end{aligned} \right\} \quad (3.7)^{**}$$

Further, making use of the first order approximation of the electromagnetic interaction, we obtain

* $|\pi\rangle$ is an abbreviation of $|\pi^+\rangle$.

** S'_F function in (3.7) still contains the electromagnetic interaction (more precisely, the interaction with the q_μ -field).

$$(\partial \mathcal{J}_{\alpha\beta; \rho\sigma}(1, 2)/\partial q_\mu(0))_0 \cong -ig_A \int d^4 x_3 (\mathcal{S}_F(1-0)\gamma_\mu \mathcal{S}_F(0-3))_{\sigma\tau} \times \mathcal{S}_{F\delta\beta}(3-2) O_{T\delta}^A O_{\rho\sigma}^A \times \dots \quad (3.8)$$

(3.5) is therefore reduced to the form

$$[1 - \mathcal{J}(1, 2)_{0|\alpha\beta; \xi\eta} \varphi_\mu(0; 1|2)_{\xi\eta}] = g_A \int d^4 x_3 (\mathcal{S}_F(1-0)\gamma_\mu \mathcal{S}_F(0-3))_{\alpha\tau} \mathcal{S}_{F\delta\beta}(3-2) O_{T\delta}^A O_{\rho\sigma}^A \varphi_\pi(3|3)_{\sigma\tau} \quad (3.9)$$

If we substitute the equation for the two-body Green function

$$[1 - \mathcal{J}(1, 2)_{0|\alpha\beta; \xi\eta} K_{\xi\eta\lambda\delta}(1, 2; 3, 4) = \mathcal{S}_{F\alpha\lambda}(1-3) \mathcal{S}_{F\delta\beta}(4-2), \quad (3.10)$$

where

$$K(1, 2; 3, 4) = \langle 0 | T(\psi^p(1) \bar{\psi}^n(2) \bar{\psi}^p(3) \psi^n(4)) | 0 \rangle,$$

into the right-hand side of (3.9) and drop $(1 - \mathcal{J}(1, 2)_{0|\alpha\beta; \xi\eta})$ from the both sides, we find the formal solution of φ_μ :

$$\varphi_\mu(0; 1|2)_{\xi\eta} = -g_A \int d^4 x_3 K_{\xi\eta\lambda\delta}(1, 2; 0, 3) (\gamma_\mu)_{\lambda\lambda'} \mathcal{S}_{F\lambda'\tau}(0-3) \times O_{T\delta}^A O_{\rho\sigma}^A \varphi_\pi(3|3)_{\sigma\tau} \quad (3.11)$$

$\langle \pi' | j_\mu(0) | \pi \rangle$ is now readily evaluated by taking the limit $(1, 2) \rightarrow (1, 2)^{\text{out}}$ or $\lim(T = \frac{1}{2}(t_1 + t_2) \rightarrow +\infty)^{14}$:

$$\langle \pi' | j_\mu(0) | \pi \rangle = g_A \int d^4 x_1 \tilde{\varphi}_{\pi'}(1|0)_{\delta\lambda} (\gamma_\mu)_{\lambda\lambda'} \mathcal{S}_{F\lambda'\tau}(0-1) O_{T\delta}^A \times \text{Tr}[O^A \varphi(1|1)] \quad (3.12)$$

with

$$\tilde{\varphi}_{\pi'}(1|0) \equiv \langle \pi' | T(\psi^p(1) \bar{\psi}^p(0)) | 0 \rangle.$$

By making use of the relation

$$(\gamma_\mu \partial_\mu + \kappa)_1 \tilde{\varphi}_{\pi'}(1|0) = g_{A'} O^{A'} \text{Tr}[O^{A'} \tilde{\varphi}_{\pi'}(1|1)] \mathcal{S}_F(1-0) \quad (3.13)$$

and separating the centre-of-mass coordinates:

$$\tilde{\varphi}_{\pi'}(x|y) = e^{-iP_\mu' X_\mu} \tilde{\tau}(z), \quad X_\mu = \frac{1}{2}(x+y)_\mu, \quad z_\mu = (x-y)_\mu,$$

the one pion expectation value of $j_\mu(0)$, $\langle j_\mu \rangle_\pi \equiv \langle \pi | j_\mu | \pi \rangle$, is finally expressed in the form

$$\langle j_\mu(0) \rangle_\pi = g_A g_{A'} N_\mu^{AA'}(P, \lambda) \tilde{F}_A(0) \Gamma_{A'}(0), \quad (3.14)$$

where

$$\left. \begin{aligned} N_\mu^{AA'}(P, \lambda) &= i/(2\pi)^4 \int \text{Tr}[\mathcal{S}_F(p) O^A \mathcal{S}_F(p+P) \gamma_\mu \mathcal{S}_F(p+P) O^{A'}] C^2(p) d^4 p, \\ \Gamma_{A'}(0) &\equiv \text{Tr}[O^{A'} \tau(0)], \\ \tilde{F}_A(0) &\equiv \text{Tr}[O^A \tilde{\tau}(0)]. \end{aligned} \right\}$$

and

N_μ depends on, besides P_μ and P^2 , the cut-off mass λ through the factor $C(p)$. $P_A(0)$ is now normalized by combining (3.14) with the condition (3.2).

The coupling constant of our $\pi-N$ interaction will be determined if we find a relation between pion field ϕ and normalized I'_A . In the rest system of a pion, the charge density ($q = -ij_4$) may be described as

$$\langle q(0) \rangle_\pi \cong 2\mu |\phi(0)|^2. \quad (\mu: \text{the pion mass}) \quad (3.15)$$

This, combined with (3.14), yields an important relation

$$2|\phi(0)|^2 \sim g_A g_{A'} \bar{N}^{AA'}(-\mu^2, \lambda) \tilde{I}'_A(0) P_{A'}(0), \quad (3.16)$$

where $\bar{N}^{AA'}$ is defined by $\bar{N}_\mu^{AA'} \equiv P_\mu N^{AA'}(P^2, \lambda)$.

As will be shown in the next section, the pseudoscalar pions are obtained when we put $O^A = \gamma_5$, in which case the relation $\tilde{I}'(0) = -I'(0)$ is found to hold. From (3.16) we see that

$$|\langle p, 0 | \mathcal{L}' | n, \pi^+ \rangle| \sim \sqrt{2} \bar{N}^{-1/2}(-\mu^2, \lambda) |\langle p | \bar{\psi}^p \gamma_5 \psi^n | n \rangle| |\phi|.$$

Therefore, the coupling constant of the $\pi-N$ interaction, G_π , may be written as

$$G_\pi \sim \bar{N}^{-1/2}(-\mu^2, \lambda) \quad (3.17)$$

in the symmetric theory. That G_π does not depend on g is a speciality of our approximation. A similar situation occurs also in the ordinary Tamm-Dancoff approach of Wentzel.¹³⁾

§ 4. Calculation of the pion mass and G_π

We are now in a position to carry out the evaluation of the pion mass and the coupling constant G_π . There are two parameters, g and λ , to be determined so that the theory may reproduce a correct pion model. Before entering into the calculation, a brief comment on the magnitude of λ would be necessary. Although it seems to be quite natural to suppose $\lambda \sim \kappa$, we want here to find out the most reasonable λ as a result, not as a starting point, of the investigation. Keeping this point in mind, we impose a weak condition on λ for the sake of facilitating the calculation. Namely, we assume $\lambda^2/\kappa^2 \gg 1$ and confine our calculation to this case. This is of course not an essential restriction to the applicability of the present theory.

Now, under this condition the differences between (2.15) and (2.15') are of little importance and the results would not significantly depend on the shape of $C(p)$.

Let us define the functions $I_k (k=0, 1, 2, 3)$ by

$$i\pi^2 \begin{bmatrix} I_0 \\ P_\mu I_1 \\ \partial_{\mu\nu} \kappa^2 I_2 + P_\mu P_\nu I_3 \end{bmatrix} \equiv \int d^4 p \frac{\begin{bmatrix} 1 \\ P_\mu \\ P_\mu P_\nu \end{bmatrix} [-\kappa^2 + \lambda^2]^2}{[p^2 + \lambda^2]^2 [(p+P)^2 + \kappa^2] [p^2 + \kappa^2]} \quad (4.1)^*$$

* $-i\epsilon$ in each factor of denominator is omitted for brevity.

and put $J_0 \equiv I_1 + I_3$ and $J_1 \equiv I_1 + I_0$, then (2.15) becomes

$$\begin{aligned} \tau(0) = & -g_r^A [J_0 (P\gamma) O^A P_A(0) (P\gamma) + I_2 \kappa^2 \gamma_\mu O^A P_A(0) \gamma_\mu \\ & + iJ_1 \kappa (P\gamma) O^A P_A(0) + iI_1 \kappa O^A P_A(0) (P\gamma) - \kappa^2 I_0 O^A P_A(0)], \end{aligned} \quad (4.2)$$

where $g_r = g/16\pi^2$. (4.2), being a covariant eigenvalue equation, is re-expressed in the rest system $P_\mu = (0, 0, 0, iP_0)$ as follows:

$$\begin{aligned} \tau(0) = & g_r^A [J_0 \mu^2 \gamma_4 O^A P_A(0) \gamma_4 - I_2 \kappa^2 \gamma_\mu O^A P_A(0) \gamma_\mu \\ & + J_1 \kappa \mu \gamma_4 O^A P_A(0) + I_1 \kappa \mu O^A P_A(0) \gamma_4 + \kappa^2 I_0 O^A P_A(0)], \end{aligned} \quad (4.3)$$

in which P_0 is to be identified with the mass μ of the composite particle. When $\tau(0)$ is expanded with respect to the sixteen Dirac matrices, $\gamma_k (k=1, 2, \dots, 16)$:

$$\tau(0) = \sum_{k=1}^{16} \chi_k \gamma_k$$

and substituted in (4.3), we obtain a set of 16 simultaneous equations for χ_k . It is easy to see that these equations are divided into four sets corresponding to the four invariant subsets of Dirac's matrices, $(1, \gamma_4)$, $(\vec{\gamma}, \gamma_4 \vec{\gamma})$, $(\gamma_5 \vec{\gamma}_4 \vec{\gamma}, \gamma_5 \vec{\gamma})$ and $(\gamma_5 \gamma_4, \gamma_5)$, giving in turn the scalar, vector, pseudovector and pseudoscalar mesons respectively. For example, the last set of $\tau(0)$:

$$\tau(0) = \langle 0 | T(\psi^n(x) \bar{\psi}^n(x)) | \rangle e^{-iP_\mu X_\mu} = \chi_{15} \gamma_5 \gamma_4 + \chi_{16} \gamma_5, \quad (4.5)$$

represents a pseudoscalar meson, because by a space reflection ψ^n and $\bar{\psi}^n$ are transformed as

$$\left. \begin{aligned} \psi & \rightarrow \psi^1 = \sum \psi \\ \bar{\psi} & \rightarrow \bar{\psi}^1 = \bar{\psi} \gamma_4 \sum^* \gamma_4 \end{aligned} \right\} \quad (\sum = i\gamma_4)$$

hence one finds

$$\begin{aligned} \tau(0) & \rightarrow \tau^1(0) = \gamma_4 \tau(0) \gamma_4 \\ & = -(\chi_{15} \gamma_5 \gamma_4 + \chi_{16} \gamma_5) = -\tau(0). \end{aligned}$$

The same arguments hold *mutatis mutandis* in other cases. If, therefore, we adopt specifically

$$-\mathcal{L}' = (g/4) \sum_{i=1}^4 : (\bar{\psi} \tau_i \gamma_5 \psi) (\bar{\psi} \tau_i \gamma_5 \psi) : \quad (4.8)$$

as the basic interaction, (4.3) has only one type of solutions which gives a pseudoscalar meson (pion).^{*} Then (4.3) reads

$$\left. \begin{aligned} \chi_{15} & = -4(J_1 - I_1) \kappa \mu \chi_{15} \\ \chi_{16} & = g_r [-4J_0 \mu^2 + 16I_2 \kappa^2 + 4I_0 \kappa^2] \chi_{16} \end{aligned} \right\} \quad (4.9)$$

^{*} It is necessary to assume that PS interactions should not exist in β -decay interaction. For the PS interaction allows the pion to decay, in a very short time ($10^{-8} \sim 10^{-9}$ sec.), into an electron and a neutrino through the interaction $p + \bar{n} \rightarrow e + \nu$. See reference 13).

with all other $\chi_k (k=1, \dots, 14)$ vanishing.

The pion mass μ is determined from the second equation of (4.9):

$$\mu^2 = J_0^{-1} [(4I_2 + I_0)\kappa^2 - (4g_r)^{-1}] \quad (4.10)$$

which is, however, not the explicit formula for μ since J_0 , I_2 and I_0 are the complicated functions of μ^2 as well as those of $\lambda^2/\kappa^2 \equiv \alpha$. A further simplification will be achieved when the condition $\alpha \gg 1$ is taken into account. Writing $I_k (k=0, \dots, 3)$ as

$$I_k(\alpha, \mu^2) = I_k^0(\alpha) + (\mu/\kappa)^2 \delta I_k(\alpha) + \dots,$$

we see that $\delta I_k(\alpha)/I_k^0(\alpha)$ tends to zero as $\alpha \rightarrow \infty$, and I_k are proportional to $\log \alpha$ for $\alpha \gg 1$ except $I_{k=2}$ which depends on α quadratically.* Hence, we can discard all δI_k except δI_2 which behaves as $\log \alpha$ for $\alpha \gg 1$. (4.10) can now be rewritten in the form

$$\mu^2 \cong (J_0')^{-1} [(4I_2^0 + I_0)\kappa^2 - (4g_r)^{-1}], \quad (4.10')$$

where

$$\left. \begin{aligned} I_2 &= I_2^0 + (\mu/\kappa)^2 \delta I_2 \\ J_0' &= J_0 - 4\delta I_2 \end{aligned} \right\} \quad (4.12)$$

(4.10') is an explicit solution of μ^2 .

In a similar way, we can neglect the μ -dependence of $\bar{N}(-\mu^2, \lambda)$ in the course of evaluating G_π . Then one finds

$$G_\pi \sim 2\pi (\log \alpha - 2)^{-1/2}. \quad (4.13)$$

The numerical results are shown in Table 1. Three cases are distinguished according to the assumed values of α , i. e., $\alpha=10$, 10^2 and 10^6 and calculations are performed by putting $(\kappa/\mu)^2=45$, $\kappa=1836 m_e$ and $G_r=g_r\lambda^2$.

Table 1

| Case | I ($\alpha=10$) | II ($\alpha=10^2$) | III ($\alpha=10^6$) |
|---------------------------------|------------------------|------------------------|------------------------|
| g (erg cm ³) | 7.83×10^{-44} | 5.79×10^{-46} | 5.48×10^{-49} |
| G_r^{-1} | 2.77 | 3.78 | ~ 4 |
| G_π | ~ 12 | ~ 4 | ~ 1.8 |
| Existence of unwanted solutions | No. | No. | No. |

To summarize the results:

(i) We may conclude that it is in fact possible to construct pions as the bound state of the nucleon-antinucleon system starting with the interaction (2.3) or (4.8), although our mathematical treatment might be rather crude.

(ii) The levels of the bound states are essentially determined by a constant G_r , reflecting the characteristic of the unrenormalizable interaction.

* See Appendix.

(iii) Reasonable values for G_π are obtained if we take $\alpha=10$ or 10^2 . The case $\alpha=10^6$, or in general too large α , may be excluded because of the smallness of corresponding G_π . Another interesting clue to the magnitude of λ will be obtained from a consideration on the spatial extension of the pion wave function $\varphi(x|y)|_{x_0=y_0}$. By taking a large straight cut-off in the three-dimensional momentum space after integrating with respect to p_0 , we find that φ is represented by a superposition of two wave functions of different range; the one is the wave function with a range λ^{-1} and the other with a range κ^{-1} . The latter, however, has a height only $\kappa\mu/\lambda^2$ times the former at the origin. Therefore, we cannot assume α to be very large.

§ 5. Inclusion of the Λ -particle and the heavy mesons

The theory so far developed in the preceding sections is no other than a covariant version of the theory of Fermi and Yang. However, it must be reminded that the compound hypothesis for elementary particles is important rather for the analysis of the nature of the heavy particles. It would, therefore, be very interesting to extend our theory to the problems of the heavy mesons and the hyperons.

A. θ - and τ -mesons

According to Sakata's idea, Λ should be regarded as the fundamental particle as well as the nucleon. Then, all heavy mesons and hyperons (except Λ itself) with some "strangeness" charge are supposed to be the composite particles including one or more Λ (or anti- Λ^*), as their constituents. In this section, however, we confine ourselves to the problem of the θ - and τ -mesons assuming both of them to be the bound states of N and $\bar{\Lambda}$, although the strangeness and the isotopic spin of τ are not yet fully determined. We need now to introduce a new basic interaction including Λ field besides (2.3):

$$-\mathcal{L}'_{\Lambda} = g_{\Lambda}^4 : [(\bar{\psi}^p O^A \psi^{\Lambda}) (\bar{\psi}^{\Lambda} O^A \psi^p) + (p \rightarrow n)] : \quad (5.1)$$

If we suppose that θ and τ are scalar (0^+) and pseudoscalar (0^-) particles respectively, and the space parity of Λ field ψ^{Λ} is odd relative to the nucleon field, we can obtain the bound states corresponding to scalar mesons from $O^A = \gamma_5$ and pseudoscalar mesons from $O^A = 1$. Hence we may leave only these two interactions in (5.1). We have thus

$$\begin{aligned} -\mathcal{L}'_{\Lambda} &= g_{\theta} : [(\bar{\psi}^p \gamma_5 \psi^{\Lambda}) (\bar{\psi}^{\Lambda} \gamma_5 \psi^p) + (p \rightarrow n)] : \\ &- g_{\tau} : [(\bar{\psi}^p \psi^{\Lambda}) (\bar{\psi}^{\Lambda} \psi^p) + (p \rightarrow n)] : \end{aligned} \quad (5.2)$$

The corresponding Nambu equation for the bound states reads

$$\left. \begin{aligned} (\gamma_{\mu} \partial_{\mu} + \kappa)_x (\gamma_{\mu}^T \partial_{\mu} - \kappa)_y \varphi^{N\bar{\Lambda}}(x|y) &= i g_{\Lambda} \{ (1 - \delta) \gamma_5 \text{Tr} [\gamma_5 \varphi^{N\bar{\Lambda}}(x|x)] \\ &- (1 + \delta) \text{Tr} \varphi^{N\bar{\Lambda}}(x|x) \} \delta(x - y), \\ (\kappa_1: \text{the mass of } \Lambda) \\ \text{with } \varphi^{N\bar{\Lambda}}(x|y) &= \langle 0 | T(\psi^N(x) \bar{\psi}^{\Lambda}(y)) | \rangle \end{aligned} \right\} \quad (5.3)$$

* Hereafter denoted by $\bar{\Lambda}$.

and g_Λ is defined by

$$g(\tilde{\tau}) = g_\Lambda(1 \pm \delta). \quad (5.4)$$

The mass of θ - or τ -mesons is obtained from (5.3) in the same way as in § 4:

$$\mu^2(\tilde{\tau}) \cong J_0'(\kappa_1)^{-1} [(4\kappa_1^2 I_2^0(\kappa_1) \mp \kappa\kappa_1 I_0(\kappa_1)) - (4g_{\Lambda\tau})^{-1}(1 \pm \delta)^{-1}], \quad (5.5)$$

where J_0' , I_2^0 and I_0 are those functions in which the argument $\alpha (= \lambda^2/\kappa^2)$ is replaced by λ^2/κ_1^2 in the corresponding functions given in Appendix.

It is well-known that the masses of θ and τ are remarkably contiguous to each other ($\sim 965 m_e$). Whether this degeneracy can be accounted for by some physical ground would be a interesting point in the theory of new particles. At first sight, the mass degeneracy may be, in our theory, realized by two alternative ways. The one is to adjust the correction factor δ in (5.4) so as to give $\mu_\theta = \mu_\tau$ and the other is to adjust g_Λ and λ so as to make $I_0(\kappa_1) = 0$ putting $\delta = 0$. In spite of these possibilities, we might as well conclude that our theory cannot explain the mass degeneracy since our explanations above seem to be quite artificial and unnatural in the point that the results depend essentially on the approximation employed which might be very crude.

The numerical results are shown in Table 2.

Table 2

| Case | I | II | III |
|---|--------|-------|----------|
| $G_{\Lambda\tau}^{-1} (= (g_{\Lambda\tau} \lambda^2)^{-1})$ | 2.57 | 3.60 | ~ 4 |
| δ^0 | -0.052 | 0.051 | ~ 0 |
| Existence of unwanted solutions | No. | No. | No. |

In the table, δ^0 is the correction factor which yields the mass degeneracy, and the "unwanted" solutions mean bound states of N and Λ .

B. Existence of singlet mesons

By making use of the total interaction Lagrangian (4.8) + (5.2), the neutral mesons are represented by a set of coupled equations of $\varphi^{p\bar{p}}$, $\varphi^{n\bar{n}}$ and $\varphi^{\Lambda\bar{\Lambda}}$. As was discussed in § 2, the neutral pion (π^0) (with isotopic spin $I=1$) is obtained as a special solution of these equations which satisfies the condition $\varphi_1^{p\bar{p}} = -\varphi_1^{n\bar{n}}$, $\varphi_1^{\Lambda\bar{\Lambda}} = 0$. On the other hand, we have the neutral meson with $I=0$, i. e., the singlet meson as a special solution satisfying $\varphi_0^{p\bar{p}} = \varphi_0^{n\bar{n}}$. That is, the singlet mesons are subjected to the equation,

$$\left. \begin{aligned} (\gamma_\mu \partial_\mu + \kappa)_x \varphi^{p\bar{p}}(x|\gamma) &= g \{ \gamma_5 \text{Tr}[\gamma_5 \varphi^{p\bar{p}}(x|x)] - 2\gamma_5 \varphi^{p\bar{p}}(x|x) \gamma_5 \} S_F(x-\gamma) \\ &+ g_\Lambda \{ (1+\delta) \varphi^{\Lambda\bar{\Lambda}}(x|x) - (1-\delta) \gamma_5 \varphi^{\Lambda\bar{\Lambda}}(x|x) \gamma_5 \} S_F(x-\gamma), \\ (\gamma_\mu \partial_\mu + \kappa_1)_x \varphi^{\Lambda\bar{\Lambda}}(x|\gamma) &= 2g_\Lambda \{ (1+\delta) \varphi^{p\bar{p}}(x|x) - (1-\delta) \gamma_5 \varphi^{p\bar{p}}(x|x) \gamma_5 \} \\ &\times S_F^\Lambda(x-\gamma), \end{aligned} \right\} \quad (5.6)$$

and the same equation for $p(\bar{p}) \rightarrow n(\bar{n})$.

The solutions are easily found in a similar way to that of §§ 3-4. Using g and g_A determined in the preceding section we obtain the following results.*

Table 3

| Type of meson | Case | I | II | III |
|---------------|------|-------------------|----|-----|
| Scalar | | No. | | |
| Vector | | $\sim 3100 m_\pi$ | | |
| Pseudovector | | No. | | |
| Pseudoscalar | | No. | | |

Therefore, only vector mesons with the mass $\sim 3100 m_\pi$ exist in the framework of the present theory.

§ 6. Discussions

In concluding the paper, we shall give a few remarks.

A. Validity of our approximation

One might suspect that the method of our approximation could not be valid, since the effective coupling strength, G_π , is not so small. But we should like to mention that the basis of validity of our treatment consists also in the assumption which requires to neglect all dynamical fluctuations (or momentum dependences) of G and the mass operator in S_F' . (See Fig. 4) If we could minimize the contributions of dynamical fluctuations by a suitable "renormalization" of κ and g under the cut-off hypothesis, our method would be convincing as far as the minimized fluctuations are small. This is the meaning of our approximation.

B. Relations to other theories of composite models

(i) The Fermi-Yang theory¹⁾

With respect to the neutron-proton interaction, (2.3) implies: $(\bar{\psi}^n O^A \psi^p) (\bar{\psi}^p O^A \psi^n)$: in which nucleon fields are ordered in the β -disintegration type. We may express our interaction as $\langle O^A \rangle_\beta$ for brevity. On the other hand, in the Fermi-Yang's theory the interaction Hamiltonian was ordered in the form: $(\bar{\psi}^n O^A \psi^p) (\bar{\psi}^p O^A \psi^n)$: (symbolically denoted by $\langle O^A \rangle_{F.Y.}$). Both types of interaction are related to each other by Fierz's identity.¹⁵⁾ In our case the identities

$$\left. \begin{aligned} \langle V \rangle_{F.Y.} &= \frac{1}{2} \langle V + A + 2P - 2S \rangle_\beta \\ \langle P \rangle_\beta &= \frac{1}{4} \langle V - T - A - P - S \rangle_{F.Y.} \end{aligned} \right\} \quad (6.1)$$

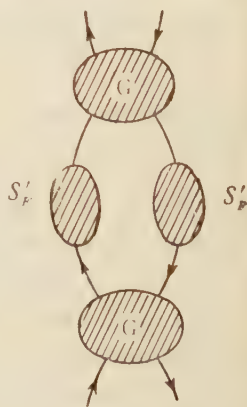


Fig. 4

* Calculations are made by putting $\delta^0=0$. This simplification substantially causes no effects

manifest most clearly a relationship between both theories. The reason why the Fermi-Yang theory contains the solutions other than the pseudoscalar one may be easily realized from the first equality of (6.1). In the case of $N-N$ interaction our coupling scheme gives

$$\langle P \rangle'_\beta = \frac{1}{4} \langle -V + T - A - P - S \rangle_{F.Y.}, \quad (6.2)$$

whereas one finds $\langle V \rangle'_{F.Y.} = -\langle V \rangle_{F.Y.}$ in the Fermi-Yang case. In our theory, the existence of $N-N$ bound state due to a short range force could be excluded on account of the inequality $\langle P \rangle'_\beta \neq \langle P \rangle_\beta$.

(ii) *The Goldhaber's model*⁽¹⁶⁾

The present method may be easily extended to the case of the Goldhaber's model where the fundamental particles from which all other heavy mesons are constructed are assumed to be N and θ instead of N and A in the Sakata model. We may suppose that the short range forces between N and $\bar{\theta}$ (anti- θ) should be derived from the direct interactions similar to (2.3). The simplest example of the interaction may be the following:*

$$\left. \begin{aligned} -\mathcal{L}' &= g : \bar{\psi} \psi : \phi^* \phi \\ \text{or} \quad -\mathcal{L}' &= g : \bar{\psi} \gamma_\mu \psi : (\partial_\mu \phi^* \cdot \phi - \phi^* \partial_\mu \phi) . \end{aligned} \right\} \quad (6.3)$$

It is worth noticing that the second interaction in (6.3) is necessary to exist since this gives different (sign reversed) interactions for $N-\theta$ and $N-\bar{\theta}$, contrary to the first one.

C. *Decays of θ and τ into leptons*

It should be noted here that a strong objection to the model for θ and τ discussed in § 5 might be raised when the β -decay interaction of A or the decay process $A \rightarrow P + e + \bar{\nu}$ is found to exist. For, in this case, one of these heavy mesons must decay into leptons with a short lifetime ($10^{-8} \sim 10^{-9}$ sec.) unless we can rule out the existence of S , P - and A -interactions from our β -decay process of A . In this connection the "two A hypothesis" proposed by Tanaka⁽¹⁾ seems to be promising. In this model the τ meson is regarded as the compound particle of N and anti- A' , the latter being a hypothetical particle which differs from A only by the parity of space reflection.

D. Thus far, we have formulated a theory of compound model for elementary particles and succeeded to some extent to explain the existence of the pion and the heavy mesons in unified manner, although our treatment might be far from completeness. To obtain the crucial test of our theory it seems to be important that the present theory be further applied to various problems such as the determination of mass levels for the Σ^+ particle and the scattering of S -wave pions by the nucleon as well as the analysis of various decay modes of the composite particles.

Acknowledgement

The author would like to express his sincere thanks to Prof. S. Sakata for his

* The isotopic spin is ignored for simplicity.

encouragement and valuable criticism to this work. Thanks are due also to Mr. S. Tanaka for his stimulating discussions.

Appendix

Evaluation of the Feynman integral (4.1)

The integrand in the right-hand side of (4.1) can be transformed into a parametric integral using the formula

$$\begin{aligned} & [p^2 + \lambda^2]^{-2} [(p+P)^2 + \kappa^2]^{-1} [p^2 + \kappa^2]^{-1} \\ &= 6 \int_0^1 \xi^2 d\xi \int_0^1 \eta (1-\eta) d\eta [(p+Q)^2 + J]^{-4}, \end{aligned} \quad (\text{A} \cdot 1)$$

where

$$\left. \begin{aligned} Q_\mu &= P_\mu \xi \eta, \\ J &= J_0 + P^2 \xi \eta (1 - \xi \eta), \\ J_0 &= \kappa^2 + (\lambda^2 - \kappa^2) \xi (1 - \eta). \end{aligned} \right\} \quad (\text{A} \cdot 2)$$

Then I_0 is expressed as follows:

$$I_0 = (\lambda^2 - \kappa^2)^2 \int_0^1 \xi^2 d\xi \int_0^1 (1-\eta) d\eta J^{-2}.$$

This may be further approximated as

$$I_0 \cong (\lambda^2 - \kappa^2)^2 \int_0^1 \xi^2 d\xi \int_0^1 (1-\eta) d\eta J_0^{-2} \quad (\text{A} \cdot 3)$$

by omitting $O(\alpha^{-1})$ (we have assumed $\alpha \gg 1$), since we can expand J^{-2} into the form

$$J^{-2} = J_0^{-2} (1 - 2P^2 J_0^{-1} \xi \eta (1 - \xi \eta) + \dots)$$

using the fact that our P^2 satisfies the relation

$$|P^2| J_0^{-1} \xi \eta (1 - \xi \eta) \leq |P^2| (4J_0)^{-1} < |P^2| (4\kappa^2)^{-1} < 1.$$

Similarly, we have

$$I_1 \cong -(\lambda^2 - \kappa^2)^2 \int_0^1 \xi^3 d\xi \int_0^1 \eta (1-\eta) d\eta J_0^{-2}. \quad (\text{A} \cdot 4)$$

I_2 and I_3 are calculated by noticing the relation

$$p_\mu p_\nu \rightarrow \frac{1}{4} p^2 \delta_{\mu\nu} + P_\mu P_\nu \xi^2 \eta^2$$

after the translation $p+Q \rightarrow p$ is performed. Thus

$$I_3 \cong (\lambda^2 - \kappa^2)^2 \int_0^1 \xi^4 d\xi \int_0^1 \eta^2 (1-\eta) d\eta J_0^{-2}. \quad (\text{A} \cdot 5)$$

On the other hand, the second term of the expansion of \mathcal{A}^{-1} :

$$\mathcal{A}^{-1} = \mathcal{A}_0^{-1} - P^2 \mathcal{A}_0^{-2} \hat{\xi} \eta (1 - \hat{\xi} \eta) + \dots$$

cannot be neglected because it gives a term comparable to other I_k 's. Hence, I_2^0 and δI_2 in (4.12) are to be defined by

$$I_2^0 = \frac{1}{2} (\lambda^2 - \kappa^2)^2 \int_0^1 \hat{\xi}^2 d\hat{\xi} \int_0^1 (1 - \eta) d\eta \mathcal{A}_0^{-1} \quad (\text{A} \cdot 6)$$

and

$$\delta I_2 = \frac{1}{2} (\lambda^2 - \kappa^2)^2 \int_0^1 \hat{\xi}^3 d\hat{\xi} \int_0^1 (1 - \hat{\xi} \eta) \eta (1 - \eta) d\eta \mathcal{A}_0^{-2} \quad (\text{A} \cdot 7)$$

respectively.

The integrations in (A.3-7) are easily performed to give

$$\left. \begin{aligned} I_0 &= \log \alpha - 2, \\ I_1 &= -\left(\frac{1}{2} \log \alpha - \frac{3}{4} \right), \\ I_3 &= \frac{1}{3} \left(\log \alpha - \frac{34}{9} \right), \\ I_2^0 &= \frac{1}{2} \left(\frac{\alpha}{2} - \log \alpha + \frac{1}{2} \right), \\ \delta I_2 &= \frac{1}{12} \left(\log \alpha + \frac{7}{6} \right), \end{aligned} \right\} \quad (\text{A} \cdot 8)$$

which are correct as far as we neglect $O(\alpha^{-1})$.

References

- 1) L. de Broglie, *Theorie des particules à spin* (1943) Paris.
- 2) T. Nakano and K. Nishijima, *Prog. Theor. Phys.* **10** (1953), 581.
K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 285.
- 3) M. Gell-Mann and A. Pais, *Proceeding Glasgow Conference*.
M. Gell-Mann, *Phys. Rev.* **92** (1953), 833.
- 4) S. Sakata, *Prog. Theor. Phys.* **16** (1956), 686.
K. Matsumoto, *Prog. Theor. Phys.* **16** (1956), 583.
S. Tanaka, *Prog. Theor. Phys.* **16** (1956), 625, 631.
- 5) C. N. Yang and E. Fermi, *Phys. Rev.* **76** (1949), 1739.
- 6) W. Heisenberg, F. Kortel and H. Mitter, *Zeits. Naturf.* **100** (1955), 425.
- 7) J. Iwadare, S. Otsuki R. Tamagaki and W. Watari, *Suppl. Prog. Theor. Phys. No. 3* (1956).
G. F. Chew and F. Low, *Phys. Rev.* **101** (1956), 1570.
- 8) S. Hayakawa, *Soryūiron Kenkyū* (Mimeographed circular in Japanese) **11** (1956), 113.
- 9) P. T. Matthews and A. Salam, *Proc. Roy. Soc.* **221** (1954), 128.
- 10) F. J. Dyson, *Phys. Rev.* **91** (1953), 1543.
- 11) Y. Nambu, *Prog. Theor. Phys.* **5** (1950), 614.
E. E. Salpeter and H. A. Bethe, *Phys. Rev.* **84** (1951), 1232.
- 12) Z. Koba, *Prog. Theor. Phys.* **5** (1950), 696.
- 13) G. Wentzel, *Phys. Rev.* **79** (1950), 710.
- 14) K. Nishijima, *Prog. Theor. Phys.* **12** (1954), 279.
K. Nishijima, *Prog. Theor. Phys.* **13** (1955), 305.
- 15) M. Fierz, *Zeits. f. Phys.* **102** (1936), 572.
- 16) M. Goldhaber, *Phys. Rev.* **101** (1956), 433.

Letters to the Editor

A New Formulation on the Electromagnetic Field

Takashi Ohmura
(formerly Kikuta)

*Department of Physics, University
of Tokyo, Tokyo*

August 31, 1956

To discuss the problem of the structure of the electron, we shall first give a formulation in the electromagnetic theory. In this formulation are introduced new scalar and pseudo-scalar fields $e(x_\mu)$ and $b(x_\mu)$ respectively, and in addition the density $\sigma(x_\mu)$ and current pseudo-vector $j(x_\mu)$ of the magnetic pole. We shall point out that the set of equations for the electromagnetic field can be combined into a single equation

$$i(\partial/\partial x_0 - \sum_{i=1}^3 \partial_i \cdot \partial/\partial x_i) \Psi = \phi, \quad (1)$$

by the use of 4×4 matrices ∂_i , together with state- and source-vectors Ψ and ϕ with four components :

$$\partial_1 = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \partial_2 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & & & 0 \end{bmatrix},$$

$$\partial_3 = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad (2)$$

$$\Psi = \begin{bmatrix} H+iE \\ ib-e \end{bmatrix}, \quad \phi = \begin{bmatrix} i+j \\ i\rho-\sigma \end{bmatrix}. \quad (3)$$

Dirac¹⁾ has developed the theory of magnetic poles, where he pointed out a possibility of quantization of charge and explained why such particles have never been observed so far. Eq. (1) has a form similar to the Dirac equation for the electron. We find the following relations among ∂_i matrices :

$$\partial_i^2 = 1, \quad \partial_i \partial_j + \partial_j \partial_i = 0, \quad (i \neq j) \quad (4)$$

$$\partial_i \partial_j = i \partial_k, \quad (i, j, k; \text{cyclic}) \quad (5)$$

It is a direct result of (4) that the field quantities satisfy the equation of harmonic oscillator with vanishing mass, when there is no charge :

$$\square \Psi \equiv (\partial^2/\partial x_0^2 - \sum \partial^2/\partial x_i^2) \Psi = 0. \quad (6)$$

Thus both the scalar and p-scalar fields $e(x_\mu)$ and $b(x_\mu)$ propagate with light velocity. These fields cause the longitudinal parts of E as well as H waves, so that it becomes possible to detect these parts by measurement of the polarization of the photon. However, as will be discussed in the subsequent letter, they can not be emitted out from electrons moving with energy less than about 100 Mev.

We can not generalize, without adding new components to Ψ , eq. (1) so as to describe a field with non-zero rest mass, because there is no matrix β commuting with all the ∂_i . This is proved simply by using (4) and (5) :

$$\begin{aligned} \beta \partial_1 &= -i \beta \partial_2 \partial_3 = i \partial_2 \beta \partial_3 \\ &= -i \partial_2 \partial_3 \beta = \partial_1 \beta = -\beta \partial_1. \end{aligned}$$

For better understanding of the physical contents of eq. (1), we will write it in the usual form :

$$\text{rot } \mathbf{H} - \partial \mathbf{E} / \partial x_0 = \mathbf{i} - \text{grad } e, \quad (7)$$

$$\text{rot } \mathbf{E} + \partial \mathbf{H} / \partial x_0 = \mathbf{j} + \text{grad } h, \quad (8)$$

$$\text{div } \mathbf{E} = \rho + \partial e / \partial x_0, \quad (9)$$

$$\text{div } \mathbf{H} = -\sigma + \partial h / \partial x_0. \quad (10)$$

As seen from (10) the magnetic field is not free from divergence, so the fields can not be derived from the usual potential (\mathbf{A}, φ). Each of \mathbf{E} and \mathbf{H} is separated into two parts $\mathbf{E}^{(1)} + \mathbf{E}^{(2)}$ and $\mathbf{H}^{(1)} + \mathbf{H}^{(2)}$: one is the solution of the equations with \mathbf{j}, σ, h zero, and other is the solution of the equations with \mathbf{i}, ρ, e zero. $\mathbf{H}^{(1)}, \mathbf{E}^{(1)}$ are expressed by the potential (\mathbf{A}, φ) and $\mathbf{H}^{(2)}, \mathbf{E}^{(2)}$ are expressed by the new potential (\mathbf{B}, ψ). Choosing a suitable gauge, the fields are then derived from a complex vector potential Φ as follows :

$$(\partial / \partial x_0 + \sum_i \partial_i \cdot \partial / \partial x_i) \Phi = i \Psi,$$

$$\Phi = \begin{bmatrix} \mathbf{A} + i \mathbf{B} \\ i \varphi - \psi \end{bmatrix}. \quad (11)$$

From (1), (4) and (11), we easily get the equation for Φ ;

$$\square \Phi = \phi. \quad (12)$$

1) P. A. M. Dirac, Phys. Rev. 74 (1948), 817.

Stability of the Electron

Takashi Ohmura

Department of Physics, University
of Tokyo, Tokyo

August 31, 1956

In the classical theory of the electron

the point charge model gives an infinite self-energy. On the other hand, if the charge is distributed over a finite region, the mass of the electron due to the electromagnetic interaction does not have correct transformation property of a particle under the Lorentz transformation. This is due to the fact that the distributed charge is not stable on account of the repulsive forces from the different parts of the same particle, and suggests that the electromagnetic field does not make up a closed system by itself. We need the "cohesive field". In the quantum electrodynamics a similar situation exists. In the present note we shall show in the framework of the classical theory that the scalar fields introduced in the preceding letter can be considered as the "cohesive field", so that the system is stable to allow the description of the correct behavior of a particle.

In order to find out the force acting on the distributed charges in the electron, we must first find out the energy-momentum tensor of the Ψ fields, $T_{\mu\nu}$. The force acting on the charges are then obtained as the 4-dimensional-divergence of this tensor. We take the following Lagrangian which leads to the wave equation in the absence of the charge :

$$L = - (1/16\pi) \sum_{\mu} [(\partial / \partial x_0 + \sum_i \partial_i \cdot \partial / \partial x_i) \Phi]_{\mu}^2 + \text{c. c.}$$

The symmetrical tensor $T_{\mu\nu}$ is constructed according to the usual procedure to give the following 4-vector force :

$$\mathbf{f} = (1/4\pi) \{ -\rho \mathbf{E}^{(1)} + \mathbf{H}^{(1)} \times \mathbf{i} - \mathbf{A} \square e - \sigma \mathbf{H}^{(2)} - \mathbf{E}^{(2)} \times \mathbf{j} + \mathbf{B} \square h \}, \quad (3)$$

$$f_4 = (i/4\pi) \{ -\mathbf{E}^{(1)} \cdot \mathbf{i} + \varphi \square e - \mathbf{H}^{(2)} \cdot \mathbf{j} + \psi \square h \}. \quad (4)$$

$T_{\mu\nu}$ and (f, f_4) are not gauge invariant, nevertheless (3) and (4) will be used to show the possibility of constructing a stable electron, because we do not know a general method of construction of $T_{\mu\nu}$ which is gauge invariant. To do this, the fields are assumed to be time-independent and spherically symmetric around the centre of the electron. To satisfy the relation $f_4=0$, it is enough to set $\rho \neq 0$, $\mathbf{j} \neq 0$, $\mathbf{i} = \sigma = 0$. Then the stability condition is,

$$f = (1/4\pi) \{-\rho E^{(1)} - B \Delta h\} = 0.$$

This equation imposes a restriction to two source functions $\rho(r)$ and $\mathbf{j}(r)$, and can be satisfied with variety of the choice of the pair (ρ, \mathbf{j}) . It will be shown below that, if source functions vanish outside the electron, h -field also vanishes outside the electron. We first notice the relation

$$\text{div } \mathbf{j} = -\Delta h. \quad (5)$$

Integrating eq. (5) with the boundary condition: $h(\mathbf{r}) \rightarrow 0$ ($|\mathbf{r}| \rightarrow \infty$), it follows that

$$h(\mathbf{r}) = (1/4\pi) \times \int \{\text{div } \mathbf{j}(\mathbf{r}') / |\mathbf{r} - \mathbf{r}'|\} d\mathbf{r}'. \quad (6)$$

Although the proper character of \mathbf{j} is not yet known, we assume that $\text{div } \mathbf{j}$ (which is spherically symmetric) is not zero only inside the electron. The integral appearing in the right side of eq. (6) has just the same form when we calculate the gravitational potential around a spherical material. Then the macroscopic conservation law: $\int \text{div } \mathbf{j}(\mathbf{r}) d\mathbf{r} = 0$ leads to the result that $h(\mathbf{r})$ vanishes outside the electron, and does not propagate far away as a wave. However, in the high energy phenomena the h -field quanta will be

emitted and will give rise to the longitudinal photon. It will be interesting to test experimentally whether the γ -ray keeps on its transverse property even in the high energy region (for instance, more than 100 Mev) as derived from the Maxwell theory or it does not as predicted from our hypothesis.

The total electric plus scalar field energy can be adjusted to be equal to the corresponding electron mass, assuming a suitable diameter for the electron.

The author would like to thank Profs. T. Yamanouchi, H. Umezawa, Drs. T. Okabayashi, and S. Tani for their interest to this work and many useful discussions concerning the related problems.

On a Composite Model for the New Particles*

Shoichi Sakata

Institute for Theoretical Physics, Nagoya University, Nagoya

September 3, 1956

Recently, Nishijima-Gell-Mann's rule¹⁾ for the systematization of new particles has achieved a great success to account for various facts obtained from the experiments with cosmic rays and with high energy accelerators. Nevertheless, it would be desirable from the theoretical standpoint

* The content of this letter was read before the annual meeting of the Japanese Physical Society held in October 1955.

A note on the same subject has also been published in Bulletin de L'académie Polonaise des Sciences (Cl. III-vol. IV, No. 6, 1956)

to find out a more profound meaning hidden behind this rule. The purpose of this work is concerned with this point.

It seems to me that the present state of the theory of new particles is very similar to that of the atomic nuclei 25 years ago. At that time, we had known a beautiful relation between the spin and the mass number of the atomic nuclei. Namely, the spin of the nucleus is always integer if the mass number is even, whereas the former is always half integer if the latter is odd. But unfortunately we could not understand the profound meaning for this even-odd rule. This fact together with other mysterious properties of the atomic nuclei, for instance the beta disintegration in which the conservation of energy seemed to be invalid, led us to a very pessimistic view-point that the quantum theory would not be applicable in the domain of the atomic nucleus. However the situation was entirely changed after the discovery of the neutron. Iwanenko and Heisenberg²⁾ proposed immediately a new model for the atomic nuclei in which neutrons and protons are considered to be

their constituents. By assuming that the neutron has the spin of one half, they explained the even-odd rule for the spins of atomic nuclei as the result of the addition law for the angular momenta of the constituents. Moreover, they could reduce all the mysterious properties of atomic nuclei to those of the neutron contained in them.

Supposing that the similar situation is realized at present, I proposed a compound hypothesis for new unstable particles to account for Nishijima-Gell-Mann's rule. In our model, the new particles are considered to be composed of four kinds of fundamental particles in the true sense, that is, nucleon, antinucleon, Λ^0 and anti- Λ^0 . If we assume that Λ^0 has such intrinsic properties as were assigned by Nishijima and Gell-Mann, we can easily get their even-odd rule for the composite particles as the result of the addition laws for the ordinary spin, the isotopic spin and the strangeness. In the next table, the models and the properties of the new particles are shown together with those of the fundamental particles in the true sense.

| Name | Model | Isotopic Spin | Strangeness | Ordinary Spin |
|----------------------------|---|---------------|-------------|---------------|
| \mathfrak{N} | | 1/2 | 0 | 1/2 |
| $\bar{\mathfrak{N}}$ | | 1/2 | 0 | 1/2 |
| Λ | | 0 | -1 | 1/2? |
| $\bar{\Lambda}$ | | 0 | 1 | 1/2? |
| π | $\mathfrak{N} + \bar{\mathfrak{N}}$ | 1 | 0 | 0 |
| $\theta(\tau)$ | $\mathfrak{N} + \bar{\Lambda}$ | 1/2 | 1 | 0? |
| $\bar{\theta}(\bar{\tau})$ | $\bar{\mathfrak{N}} + \Lambda$ | 1/2 | -1 | 0? |
| Σ | $\mathfrak{N} + \bar{\mathfrak{N}} + \Lambda$ | 1 | -1 | 1/2? |
| Ξ | $\bar{\mathfrak{N}} + \Lambda + \Lambda$ | 1/2 | -2 | 1/2? |

Here \mathfrak{N} and $\bar{\mathfrak{N}}$ denote nucleon and antinucleon respectively, whereas Λ and $\bar{\Lambda}$ denote Λ^0 and anti- Λ^0 respectively³⁾.

So far as the internal structure is not concerned, our model for new particles is identical with that of Nishijima and Gell-

Mann. However, it should be stressed that the curious properties of the new particles could be reduced to those of A^0 , just like the mysterious properties of the atomic nuclei were reduced to those of neutron. Hence our theory contains less arbitrary elements than was the case for original one of Nishijima and Gell-Mann.

Though the rigorous treatment of our model is a very hard task⁴⁾, it is worthwhile to notice that most of the composite particles which seem to be stable against the strong interaction can be identified with the well-known new particles, and that there are possibilities of predicting some more new particles which have not been discovered up till now.⁵⁾

Finally, it should be remarked that there are some other arguments in favour of the compound hypothesis for the elementary particles. In spite of the great success achieved by the advent of Tomonaga-Schwinger's technique, it has recently become clear that we could not avoid the internal inconsistency of the quantum field theory, so far as the point model for elementary particles was adopted. Moreover, in the case of π -meson, the cut-off prescription has recently been proved to be very powerful in order to account for the experimental results. These facts indicate strongly the necessity of substantial innovations in the model for the elementary particles, though some change has already been made by the discovery of the renormalization technique. Landau pointed out that the model for the electron would possibly be changed by the effect of the gravitational field. But in the case of π -meson we must look for another effect, because the cut-off radius is found to be as large as the order of the nucleon Com-

pton wave length in contrast to $e^0/mc^0 \cdot e^{-137} \sim 10^{-58}$ cm which appeared in the quantum electrodynamics.⁶⁾

- 1) T. Nakano and K. Nishijima, *Prog. Theor. Phys.* **10** (1953), 581; K. Nishijima, *Prog. Theor. Phys.* **12** (1954), 107; **13** (1955), 285; M. Gell-Mann, *Phys. Rev.* **92** (1953), 833.
- 2) D. Iwanenko, *Nature* **129** (1932), 798; W. Heisenberg, *ZS. Phys.* **77** (1932), 1.
- 3) Markov (*Rep. Acad. Sci. USSR*, 1955) proposed also a composite model which is very similar to ours. It should be remarked that our model may be considered as a generalization of the π -meson model proposed by Fermi and Yang (*Phys. Rev.* **76** (1948), 1739), and that it will throw a new light on Heisenberg's theory of elementary particles (*Zs. Naturf.* **9a** (1954), 291; **10a** (1955), 425), in which only one kind of "Urmaterie" is assumed.
- 4) S. Tanaka, *Prog. Theor. Phys.* **16** (1956), 625.
- 5) Z. Maki, *Prog. Theor. Phys.* **16** (1956), 667.
- 6) K. Matsumoto, *Prog. Theor. Phys.* **16** (1956), 583.
- 7) M. A. Markov, *Uspekhi Fiz. Nauk* **51** (1953), 317; L. Landau et al., *DAN.* **95** (1954), 497, 733, 1177; **96** (1954), 261; **102** (1955), 489; S. Kamefuchi & H. Umezawa, *Prog. Theor. Phys.* **15** (1956), 298; *Nuovo Cimento* **3** (1956), 1060.

Special Theory of Relativity and the Structure of Elementary Particles

Hideki Yukawa

*Research Institute for Fundamental Physics,
Kyoto University, Kyoto*

October 19, 1956

The present field theory of elementary particles is based on both quantum mechanics and the special theory of relativity. Since we have for so long been unsuccessful

in constructing a unified and consistent theory, we are beginning to doubt that the rigid framework of the special theory of relativity and quantum mechanics will remain without modification in the future theory of elementary particles. For this reason, it is all the more important at the present time to investigate whether the space-time framework of the special theory of relativity is appropriate for discussions of the possible internal structure of elementary particles. The so-called non-local field theory was in fact an attempt to extend the space-time concept of special relativity to the interior of the elementary particles.¹⁾ Thus, apart from the detailed procedure of formulation, it may be worthwhile to discuss in more general terms the extent to which the idea of internal motion in non-local field theory is useful in the search for the deeper meaning of the new degrees of freedom which were introduced recently for the purpose of systematizing the strongly interacting particles.

Let us suppose that the field for each type of particle is a function not only of one set of space-time coordinated x_μ , which are related to the motion of the particle as a whole, but also of variables which are related, in a broad sense, to the structure of the particle. Suppose further that some of these variables, ξ_μ , could be identified with internal coordinates designating a point in the internal space-time world of the particle which we suppose to have the same structure as the external world. Thus, the field can be denoted by $\psi(x_\mu, \xi_\mu, \dots)$, where ψ may consist of a number of components.

Now, let us introduce the following three assumptions in order to make the situation as clear and simple as possible:

I. In the zero approximation, the field for each of the strongly interacting particles can be written as a product of external and internal wave function:

$$\psi(x_\mu, \xi_\mu, \dots) = \psi_{\text{ext}}(x_\mu) \psi_{\text{int}}(\xi_\mu, \dots).$$

Both ψ_{ext} and ψ_{int} may have a number of components, which transform linearly under a Lorentz transformation. Thus, by considering the Lorentz transformations of x_μ and ξ_μ separately as well as the simultaneous transformations of both, the external spin and parity can be discriminated from the internal as well as from the total spin and parity.

II. In the case of strong interaction, each of the interaction terms in the Lagrangian or the Hamiltonian for a system of elementary particles should be constructed from three fields in such a way that it is a relativistically invariant product of two factors, of which one is an invariant product of three external wave functions and the other is an invariant product of corresponding internal wave functions.

III. In the case of weak interaction on the other hand, the interaction term in the Lagrangian should again be relativistically invariant as a whole, but need not necessarily be so for the external and internal parts separately. For instance, it could be a scalar product of external and internal spinors, vectors, etc. In the following, we shall consider only the case of a scalar product of two spinor functions.

If, in addition to these three assumptions, we take account of the charge independence of the pion-nucleon interaction, we obtain the following scheme as one of the simplest possible examples of assignment of external and internal wave functions to each type of strongly inter-

acting particle :

| | ϕ_{ext} | ϕ_{int} | total spin (and parity) |
|-----------|---------------------|---------------------|-------------------------|
| π | pseudoscalar | scalar | 0^- |
| N | Dirac | scalar | $1/2$ |
| Λ | scalar | Dirac | $1/2$ |
| Σ | scalar | Dirac | $1/2$ |
| Ξ | Dirac | vector | $1/2, 3/2$ |
| K | Dirac | anti-Dirac | $0^\pm, 1^\pm$ |

Here anti-Dirac means that the wave function transforms like the charge conjugate of a Dirac wave function. If we adopt this scheme, it follows from the second assumption that the strong interactions are restricted essentially to cases such as $\bar{N}\pi N$, $\bar{\Lambda}KN$, $\bar{\Sigma}KN$, $\bar{K}\pi K$, and $\bar{\Xi}\pi\Xi$. Similarly, according to the third assumption, the weak interactions could be restricted to the cases $N\pi\Lambda$, $\bar{N}\pi\Sigma$, $\bar{\Lambda}\pi\Xi$, $\bar{\Sigma}\pi\Xi$, and possibly $\pi\pi K$. All of these strong and weak interactions are acceptable, and some of them at least are necessary in order to account qualitatively for both the production and decay processes of strongly interacting particles. The total spin and parity are conserved always, but the weak interaction is distinct from the strong interaction in that the former gives rise to the exchange of angular momentum of an amount $\hbar/2$ between the external and internal degrees of freedom of the interacting particles.

If we compare the above scheme with the Gell-Mann-Nishijima scheme,²⁾ we notice immediately that the strangeness quantum number is equal to twice the internal spin. The isospin quantum number are purposely left out of the above scheme because the isospin space cannot be easily

connected with ordinary space-time owing to the peculiarity of the electromagnetic interaction.

Perhaps the most serious defect of the above scheme is that, in spite of the introduction of internal spin and parity, we cannot resolve the wellknown dilemma of the 2π and 3π decays of K -mesons unless we adopt either a more complicated scheme involving higher spins or simply give up the conservation of total parity. It may, however, be premature to say anything very definite about this crucial problem.

This work was mostly done during the author's stay at the Radiation Laboratory of the University of California in Berkeley. The author is grateful for the hospitality of Professor E. Larwrence and for stimulating discussions with the members of Theoretical Physics Group of the Radiation Laboratory. His thanks are due also to Professors Z. Koba and D. Yennie for instructive comments.

- 1) H. Yukawa, Phys. Rev. **77** (1950), 219; **80** (1950), 1047; **91** (1953), 415, 416; Science, **121** (1955), 405.
O. Hara, T. Marumori, Y. Ohnuki and H. Shimodaira, Prog. Theor. Phys. **12** (1954), 177.
J. Rayski, Nuovo Cimento **10** (1953), 1729; **12** (1954), 815.
M. A. Markov, Uspekhi Fiz. Nauk **51** (1953), 317.
- 2) K. Nishijima, Prog. Theor. Phys. **12** (1954), 107; **13** (1955), 304.
M. Gell-Mann, Phys. Rev. **92** (1953), 833.
M. Gell-Mann and A. Pais, Proceedings of the Glasgow Conference (1954), 342.

AUTHORS INDEX TO VOLUME 16

Araki-G., Partial Description of Quantum-Mechanical System 197
Araki-G., Free-Electron Model of Benzene .. 212
Araki-G., Free-Electron Ring Model of Polyacenes 265
Falkoff-D., Integral over Path Formulation of Statistical Theory of Irreversible Processes ..530(L)
Fujii-Y (see Hara-O).....149(L), 150(L)
Fujimoto-Y, Fukuzawa-F & Okai-S, Some Remarks on the Nuclear Level Spacing ..246(L)
Fujita-J., On Quantum-Mechanical Nuclear Dipole Vibrations 112
Fukuchi-M The Energy Band Structure of the Metallic Copper..... 222
Fukuzawa-F (see Fujimoto-Y) 246(L)
Furuichi-S, Kodama-T, Sugahara-Y, Yonezawa-M& Wakasa-A, The Interaction of $K_{\mu 3}$ and K_{e3} and Fermi Interaction..... 64(L)
Ghosh-N. N., On a Solution of Field Equations in Einstein's Unified Field Theory, I.....421
Hara-O & Fujii-Y, An Attempt of Generalizing the Hypothesis of Charge Independence....
..... 149(L)
Hara-O & Fujii-Y, A Conjecture on the Possible Existence of New Hyperons150(L)
Hatano-S, Note on the Decay of Ξ Particles 248(L)
Hayakawa-S & Kitao-K, Energy Loss of a Charged Particle Traversing Superconductors. 131
Hayakawa-S & Kitao-K, Energy Loss of a Charged Particle Traversing Ionized Gas and Injection Energies of Cosmic Rays..... 139
Hayakawa-S, Hayashi-C, Imoto-M & Kikuchi-K, Helium Capturing Reactions in Stars 507
Hayashi-C (see Hayakawa-S) 507
Hayashi-C & Nishida-M, Formation of Light Nuclei in the Expanding Universe..... 613
Higashi-S, Oda-M, Oshio-S, Shibata-H, Watanabe-K & Watase-Y, On the Energy Dependence of the Cross Section for the Production of the Penetrating Shower Underground..250(L)
Honda-N (see Matsubara-T) 447
Hori-S, On the Determination of the Spins of Λ^0 and θ^0 Particles 189
Hsieh-S. H., Nature of Nuclear Forces indicated by the Photodisintegration of the Deuteron, II 68(L)
Ikeda-K, On the Theory of Condensation .. 361

Ikeda-M & Miyachi-Y, On the Extended Framework for the Description of Elementary Particles 537
Imoto-M (see Hayakawa-S) 507
Iso-C & Kawaguchi-M, Note on the Decay Interactions of Hyperon and Heavy Mesons 177
Iso-C & Iwao-S, On the Effects of Heavy Unstable Particles to the Anomalous Magnetic Moments of Nucleons 410(L)
Ito-D & Minami-S, Meson-Nucleon Scattering in Low Energy Region..... 548
Iwadare-J, Otsuki-S, Sano-M, Takaji-S, Tamagaki-R, & Watari-W, Two-Nucleon Problem with Pion Theoretical Potential, I, II, III, IV 455, 472, 604, 658
Iwao-S (see Iso-S) 410(L)
Kajikawa-R, Sasakawa-T & Watari-W, Inelastic Scattering of Nucleons by Nuclei..152(L)
Kanazawa-A & Sugawara-M, Nucleon Magnetic Moments in Ps-Ps Intermediate Coupling Meson Theory 95
Kasuya-T, A Theory of Metallic Ferro- and Antiferromagnetism on Zener's Model..... 45
Kasuya-T, Electrical Resistance of Ferromagnetic Metals 58
Katsura-S, On the Bose-Einstein Condensation 589
Kawaguchi-M (see Iso-C) 177
Kikuchi-K (see Hayakawa-S) 507
Kimura-T, On the Quantum Field Theory of the Interaction between the Graviton and the Matter Field, I, II 157, 555
Kitamura-T & Oda-M, On the Nuclear Interaction of Meson below Ground..... 252(L)
Kitao-K (see Hayakawa-S) 131, 139
Klemens-P. G., On the Theory of Thermal Conductivity of Pure Metals at Low Temperatures 154(L)
Kodama-T (see Furuichi-S)..... 64(L)
Konuma-M, The Range of the Interaction between Different Fermions 259
Maki-Z, On a Theory of the Composite Model of Elementary Particles..... 667
Marumori-T, Suekane-S & Yamamoto-A, Nuclear Deformability and Shell Model.... 320
Matsubara-T, & Matsuda-H, A Lattice Model of Liquid Helium, I 409(L)
Matsubara-T, Morita-A & Honda-N, Theory of Bose-Einstein Condensation of an Imperfect

- Bose-Einstein Gas 447
Matsubara-T & Matsuda-H, A Lattice Model of Liquid Helium, II 569
Matsuda-H (see Yamamoto-T) 269
Matsuda-H (see Matsubara-T) 569
Matumoto-K, Some Consequences of the Compound Hypothesis for Elementary Particles.. 583
Minami-S, Interaction of K-Meson and Nucleon 66(L)
Minami-S (see Ito-D) 548
Miyachi-Y (see Ikeda-M) 537
Morita-A (see Matsubara-T) 447
Moriya-T, Nuclear Magnetic Relaxation in Antiferromagnetics, I, II 23, 641
Munakata-Y (see Sugano-R) 532(L)
Murota-T, Ueda-A & Tanaka-H, The Creation of an Electron Pair by a Fast Charged Particle 482
Murota-T & Ueda-A, On the Foundation and the Applicability of Williams-Weizsäcker Method 497

Nagasaki-M, Deuteron Stripping and the Compound Nucleus 429
Nakagawa-K, Ohmura-T, Takebe-H & Obi-S, Nuclear Reactions in the Later Stage of the Stellar Evolution 341
Namiki-M (see Saito-N) 71
Nanda-V.S & Trikha-S.K, Thermodynamic Properties of He^3 254(L)
Nishida-M (see Hayashi-C) 613
Nishiyama-T, State Vectors in the Collective Description 244(L)

Obi-S (see Nakagawa-K) 341
Oda-M (see Higashi-S) 250(L)
Oda-M (see Kitamura-T) 252(L)
Ohmura-T, Extensions of Variational Method, III 231
Ohmura-T (see Nakagawa-K) 341
Ohmura-T, A New Formulation on the Electromagnetic Field 684(L)
Ohmura-T, Stability of the Electron 685(L)
Okai-S (see Fujimoto-Y) 246(L)
Oshio-S (see Higashi-S) 250(L)
Otsuki-S (see Iwadare-J) 455, 472, 604, 658
Ouchi-T & Senba-K, Charge Independence Hypothesis and Leptons 528(L)

Saito-N & Namiki-M, On the Quantum Mechanics-like Description of the Theories of the Brownian Motion and Quantum Statistical Mechanics..... 711
Sakata-S, On a Composite Model for the New Particles 686(L)
Sano-M (see Iwadare-J) 658(L)
Sasakawa-T (see Kajikawa-R) 152(L)
Sasakawa-T, On the Energy Matrix of the Tensor Forces 256(L)
Senba-K (see Ouchi-T) 528(L)
Shibata-H (see Higashi-S) 250(L)
Shimazu-H, A Relativistic Wave Equation for a Particle with Two Mass States of Spin 1 and 0 287
Suekane-S (see Marumori-T) 320
Sugahara-Y (see Furuichi-S) 64(L)
Sugano-R & Munakata-Y, On the Bethe-Salpeter Equation for the Scalar-Spinor Particle System 532(L)
Sugawara-M (see Kanazawa-A) 95

Takagi-S (see Iwadare-J) 658
Takebe-H (see Nakagawa-K) 341
Tamagaki-R (see Iwadare-J) 455, 472, 604,
Tanaka-H (see Murota-T) 482
Tanaka-S, The Composite Model for New Unstable Particles, I, II 625, 631,
Trikha-S.K (see Nanda-V.S) 254(L)

Ueda-A (see Murota-T) 482, 497
Ui-H, General Theory of Deuteron Induced Reaction, I 299

Wakasa-A (see Furuichi) 64(L)
Watanabe-K (see Higashi-S) 250(L)
Watanabe-Y, A Classical Theory of the Collective Description 1
Watanabe-Y, Constraint in a Quantum System 534(L)
Watari-W (see Kajikawa-R) 152(L)
Watari-W (see Iwadare-J) 455, 472, 604, 658
Watase-Y (see Higashi-S) 250(L)

Yamamoto-A (see Marumori-T) 320
Yamamoto-T & Matsuda-H, On the Grand Canonical Distribution Method of Statistical Mechanics..... 269
Yonezawa-M (see Furuichi-S) 64(L)
Yukawa-H, Special Theory of Relativity and the Structure of Elementary Particles..... 688(L)

